

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: June 16, 2004, 11:08:12 ; Search time 49 Seconds
(without alignments)
1586.865 Million cell updates/sec

Title: US-10-018-672-2

Sequence: 1 MAFKINGICALAGIALAG.....TDEVEAKKQFDGVIRKM 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1158786 seqs, 281726120 residues

Total number of hits satisfying chosen parameters: 1158786

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Published Applications RA:
2: /cgn2_6/ptodata/2/pubppa/PCR_NEM_PUB.pdb:
3: /cgn2_6/ptodata/2/pubppa/US06_NEM_PUB.pdb:
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pdb:
5: /cgn2_6/ptodata/2/pubppa/US07_NEM_PUB.pdb:
6: /cgn2_6/ptodata/2/pubppa/PCR_NEM_PUBCOMB.pdb:
7: /cgn2_6/ptodata/2/pubppa/US08_NEM_PUB.pdb:
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pdb:
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pdb:
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pdb:
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pdb:
12: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pdb:
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pdb:
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pdb:
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pdb:
16: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pdb:
17: /cgn2_6/ptodata/2/pubppa/US60_NEM_PUB.pdb:
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pdb:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1368	98.9	276	US-10-282-122A-63386	Sequence 63386, A
2	1007.5	72.8	273	US-10-380-817-2	Sequence 2, Appl1
3	1007.5	72.8	273	US-10-380-817-4	Sequence 4, Appl1
4	1007.5	72.8	273	US-10-380-817-8	Sequence 8, Appl1
5	1007.5	72.8	273	US-10-380-817-10	Sequence 10, Appl1
6	1006.5	72.8	273	US-09-815-242-11067	Sequence 11067, A
7	1006.5	72.8	273	US-10-282-122A-58240	Sequence 58240, A
8	1006.5	72.8	273	US-10-380-817-6	Sequence 6, Appl1
9	888	64.2	276	US-10-282-122A-67385	Sequence 67385, A
10	793.5	57.4	271	US-10-282-122A-76195	Sequence 76195, A
11	782.5	56.6	271	US-09-815-242-10050	Sequence 10050, A
12	782.5	56.6	271	US-10-282-122A-43291	Sequence 43291, A
13	780.5	55.4	271	US-10-282-122A-55826	Sequence 55826, A
14	775.5	54.1	271	US-10-282-122A-59600	Sequence 59600, A
15	757.5	54.8	271	US-10-282-122A-68391	Sequence 68391, A

16	757.5	54.8	271	US-10-282-122A-78174	Sequence 78174, A
17	751.5	54.3	275	US-10-282-122A-77174	Sequence 77174, A
18	743.5	53.8	240	US-09-815-242-11656	Sequence 11656, A
19	562	40.6	295	US-10-282-122A-47704	Sequence 47704, A
20	551	39.8	262	US-10-282-122A-51371	Sequence 51371, A
21	549.5	39.7	272	US-10-282-122A-50375	Sequence 50375, A
22	545	39.4	256	US-10-282-122A-61127	Sequence 61127, A
23	540	39.0	268	US-10-282-122A-49256	Sequence 49256, A
24	527.5	38.1	270	US-10-282-122A-49339	Sequence 49339, A
25	526.5	38.6	265	US-10-282-122A-51201	Sequence 51201, A
26	506.5	36.6	261	US-10-282-122A-67692	Sequence 67692, A
27	502.5	36.3	260	US-09-815-242-12110	Sequence 12110, A
28	502.5	36.3	260	US-10-282-122A-66844	Sequence 66844, A
29	497.5	36.0	259	US-09-815-242-5146	Sequence 5146, Ap
30	494	35.7	263	US-10-282-122A-43469	Sequence 43469, A
31	488	35.3	256	US-10-282-122A-52976	Sequence 52976, A
32	487	35.2	257	US-10-282-122A-67627	Sequence 67627, A
33	481	34.8	277	US-10-282-122A-69359	Sequence 69359, A
34	481	34.8	277	US-10-282-122A-57209	Sequence 57209, A
35	477.5	34.5	257	US-10-282-122A-54418	Sequence 54418, A
36	472.5	34.2	270	US-10-282-122A-54496	Sequence 54496, A
37	472.5	34.1	270	US-10-282-122A-54838	Sequence 52838, A
38	468	33.8	256	US-10-282-122A-54417	Sequence 54417, A
39	467.5	33.8	271	US-09-815-242-11460	Sequence 11460, A
40	467.5	33.8	271	US-10-282-122A-59034	Sequence 59034, A
41	467.5	33.8	271	US-09-815-242-11624	Sequence 11624, A
42	462.5	33.4	271	US-10-335-977-4882	Sequence 4882, Ap
43	462.5	33.4	271	US-10-335-977-4883	Sequence 4883, Ap
44	462.5	33.4	271	US-10-282-122A-60558	Sequence 60558, A
45	462.5	33.4	273	US-10-282-122A-60558	Sequence 60558, A

RESULT 1
US-10-282-122A-63386
Sequence 63386, A
Publication No. US20040029123A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haeselbeck, Robert

APPLICANT: Ohlson, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: EPIIIRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

ALIGNMENTS
102 (e)
Qual 6/25 101

;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 63386
;; LENGTH: 276
;; TYPE: PRF
;; ORGANISM: Moraxella catarrhalis
US-10-282-122A-63386

Query Match 98.9%; Score 1368; DB 12; Length 276;
Best Local Similarity 99.3%; Pred. No. 1.8e-116;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAAISTKTAQTIKVGWAGPQAAEVAAGVAK 60
DB 1 MNFGKIFGICALASGIALAGCSNOSNEPAAISDPAQTIKVGWAGPQAAEVAAGVAK 60
QY 61 EKNITVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYVP 120
DB 61 EKNITVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYVP 120
QY 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180
DB 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180
QY 181 PKLVIKEVDTSVAAARALDVIDLAVVNNNTYAGVGLTASENGVEVEDKDSFYNNITVARA 240
DB 181 PKLVIKEVDTSVAAARALDVIDLAVVNNNTYAGVGLTASENGVEVEDKDSFYNNITVARA 240
QY 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQKGVIKGW 276
DB 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQKGVIKGW 276

RESULT 2

US-10-380-817-2
;; Sequence 2, Application US/10380817
;; Publication No. US20040039169A1
;; GENERAL INFORMATION:
;; APPLICANT: Thonard, Joelle
;; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
;; FILE REFERENCE: BM45419
;; CURRENT APPLICATION NUMBER: US/10/380,817
;; CURRENT FILING DATE: 2003-03-18
;; PRIOR APPLICATION NUMBER: PCT/EP01/10979
;; PRIOR FILING DATE: 2001-09-18
;; PRIOR APPLICATION NUMBER: GB 002292.2
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 2
;; LENGTH: 273
;; TYPE: PRF
;; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-2

Query Match 72.8%; Score 1007.5; DB 12; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.4e-83;
Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAAISTKTAQTIKVGWAGPQAAEVAAGVAK 60
DB 1 MNFGKIFGICALASGIALAGCSNOSNEPAAISDPAQTIKVGWAGPQAAEVAAGVAK 57
QY 61 EKNITVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYVP 120
DB 61 EKNITVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYVP 117
QY 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180

DB 118 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 177
QY 181 PKLVIKEVDTSVAAARALDVIDLAVVNNNTYAGVGLTASENGVEVEDKDSFYNNITVARA 240
DB 178 PKLVIKEVDTSVAAARALDVIDLAVVNNNTYAGVGLTASENGVEVEDKDSFYNNITVARA 237
QY 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQKGVIKGW 276
DB 238 DNKSKAIQDFVKAAYQTDVEAEAKKQFQKGVIKGW 273

RESULT 3

US-10-380-817-4
;; Sequence 4, Application US/10380817
;; Publication No. US20040039169A1
;; GENERAL INFORMATION:
;; APPLICANT: Thonard, Joelle
;; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
;; FILE REFERENCE: BM45419
;; CURRENT APPLICATION NUMBER: US/10/380,817
;; CURRENT FILING DATE: 2003-03-18
;; PRIOR APPLICATION NUMBER: PCT/EP01/10979
;; PRIOR FILING DATE: 2001-09-18
;; PRIOR APPLICATION NUMBER: GB 002292.2
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 4
;; LENGTH: 273
;; TYPE: PRF
;; ORGANISM: No. US20040039169A1-typeable Haemophilus Influenzae
US-10-380-817-4

Query Match 72.8%; Score 1007.5; DB 12; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.4e-83;
Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAAISTKTAQTIKVGWAGPQAAEVAAGVAK 60
DB 1 MNFGKIFGICALASGIALAGCSNOSNEPAAISDPAQTIKVGWAGPQAAEVAAGVAK 57
QY 61 EKNITVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYVP 120
DB 61 EKNITVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLTIVGNTFYVP 117
QY 121 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 180
DB 118 LAGSTKIKTLNELKDGTIAVPNDPSLALILLEKQGLIKKXNTNLFSTLTDIVEN 177
QY 181 PKLVIKEVDTSVAAARALDVIDLAVVNNNTYAGVGLTASENGVEVEDKDSFYNNITVARA 240
DB 178 PKLVIKEVDTSVAAARALDVIDLAVVNNNTYAGVGLTASENGVEVEDKDSFYNNITVARA 237
QY 241 DNKSKAIQDFVKAAYQTDVEAEAKKQFQKGVIKGW 276
DB 238 DNKSKAIQDFVKAAYQTDVEAEAKKQFQKGVIKGW 273

RESULT 4

US-10-380-817-8
;; Sequence 8, Application US/10380817
;; Publication No. US20040039169A1
;; GENERAL INFORMATION:
;; APPLICANT: Thonard, Joelle
;; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASE202
;; FILE REFERENCE: BM45419
;; CURRENT APPLICATION NUMBER: US/10/380,817
;; CURRENT FILING DATE: 2003-03-18
;; PRIOR APPLICATION NUMBER: PCT/EP01/10979
;; PRIOR FILING DATE: 2001-09-18

PRIOR APPLICATION NUMBER: GB 0022992.2
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 273
TYPE: PRT
ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-8

Query Match 72.8%; Score 1007.5; DB 12; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.4e-83;
Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAISTKTAOTIKYGVAGPEQAVAEVAGVAK 60
DB 1 MKLQOLFALTAIALSLVLTGCKEDKKPEAA--AAPLKIKGVWMSGPHQVAEIAKAK 57
QY 61 EKYNTLVELFENDYAMENSAVSKGELDANAMQHKPYLEKSOEKGNNLVYVNTFYVP 120
DB 58 EKYGLDVQFVEFENDYALPNEAVSKGDLNANMOKHPEYLEDKAKANNLVYVNTFYVP 117
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKKDKNTLFFSTLDIYEN 180
DB 118 LAGYSKKIKNNVELDGAQKVVVPNDPTNRRALILIEKQGLIKKDKANNLSTVLDIYEN 177
QY 181 PKKLVKEVDSVAARAIDVDVLAANNVYAGVGLTSENGVFEEDKDSFYNNIIVARA 240
DB 178 PKKLNTTEVDSVAARALDDVDLAANNVYAGVGLNADDDGVFEEDKDSFYNNIIVARS 237
QY 241 DNKDSKAIODFYKAYOTDEVEAEAKKQFQDGVYKGM 276
DB 238 DNKDSKAVODFYKSYOTEVEVYQEAQKHFDGVYKGM 273

RESULT 5
US-10-380-817-10
Sequence 10, Application US/10380817
Publication No. US20040039169A1
GENERAL INFORMATION:
APPLICANT: Thonard, Joselle
TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASR202
FILE REFERENCE: BM45419
CURRENT FILING DATE: 2003-03-18
CURRENT FILING DATE: 2003-03-18
PRIOR APPLICATION NUMBER: PCT/EP01/10979
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: GB 0022992.2
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 273
TYPE: PRT
ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-10

Query Match 72.8%; Score 1007.5; DB 12; Length 273;
Best Local Similarity 71.7%; Pred. No. 1.4e-83;
Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAISTKTAOTIKYGVAGPEQAVAEVAGVAK 60
DB 1 MKLQOLFALTAIALSLVLTGCKEDKKPEAA--AAPLKIKGVWMSGPHQVAEIAKAK 57
QY 61 EKYNTLVELFENDYAMENSAVSKGELDANAMQHKPYLEKSOEKGNNLVYVNTFYVP 120
DB 58 EKYGLDVQFVEFENDYALPNEAVSKGDLNANMOKHPEYLEDKAKANNLVYVNTFYVP 117
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKKDKNTLFFSTLDIYEN 180
DB 118 LAGYSKKIKNNVELDGAQKVVVPNDPTNRRALILIEKQGLIKKDKANNLSTVLDIYEN 177

QY 181 PKKLVKEVDSVAARAIDVDVLAANNVYAGVGLTSENGVFEEDKDSFYNNIIVARA 240
DB 178 PKKLNTTEVDSVAARALDDVDLAANNVYAGVGLNADDDGVFEEDKDSFYNNIIVARS 237
QY 241 DNKDSKAIODFYKAYOTDEVEAEAKKQFQDGVYKGM 276
DB 238 DNKDSKAVODFYKSYOTEVEVYQEAQKHFDGVYKGM 273

RESULT 6
US-09-815-242-11067
Sequence 11067, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA-011A
CURRENT FILING DATE: 2001-03-21
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11067
LENGTH: 273
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11067

Query Match 72.8%; Score 1006.5; DB 9; Length 273;
Best Local Similarity 71.4%; Pred. No. 1.7e-83;
Matches 197; Conservative 28; Mismatches 48; Indels 3; Gaps 1;

QY 1 MNFGKINGICALASGIALAGCSNOSNEPAISTKTAOTIKYGVAGPEQAVAEVAGVAK 60
DB 1 MKLQOLFALTAIALSLVLTGCKEDKKPEAA--AAPLKIKGVWMSGPHQVAEIAKAK 57
QY 61 EKYNTLVELFENDYAMENSAVSKGELDANAMQHKPYLEKSOEKGNNLVYVNTFYVP 120
DB 58 EKYGLDVQFVEFENDYALPNEAVSKGDLNANMOKHPEYLEDKAKANNLVYVNTFYVP 117
QY 121 LAGYSTKIKTLNELKDGATIAVPNDPSNLARALILIEKQGLIKKDKNTLFFSTLDIYEN 180
DB 118 LAGYSKKIKNNVELDGAQKVVVPNDPTNRRALILIEKQGLIKKDKANNLSTVLDIYEN 177
QY 181 PKKLVKEVDSVAARAIDVDVLAANNVYAGVGLTSENGVFEEDKDSFYNNIIVARA 240
DB 178 PKKLNTTEVDSVAARALDDVDLAANNVYAGVGLNADDDGVFEEDKDSFYNNIIVARS 237
QY 241 DNKDSKAIODFYKAYOTDEVEAEAKKQFQDGVYKGM 276
DB 238 DNKDSKAVODFYKSYOTEVEVYQEAQKHFDGVYKGM 273

```

RESULT 7
US-10-282-122A-58240
; Sequence 58240, Application US/10282122A
; Publication NO. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/1191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/256,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/250,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58240
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-10-282-122A-58240

```

[illegible][illegible]

		72.8% ; Score 1006.5; DB 12; Length 273;
		Best Local Similarity 71.7% ; Pred. No. 17e-83;
		Matches 188; Conservative 26; Mismatches 49; Indels 3; Gaps 1
QY	1	KNFGKINGICALAGCAGCCNSQNSNPAAIISTKAQTIVKGVMAQPEQAVAEVAGQAK 60
Db	1	MLKQLFAITIASALVLTGCKEDKPEEA---AAPLKIYGVMSGPEHQVAIEIAKAK 57
QY	61	EKYNTLVEJVEFNDPAMNSAVSKGELDAANAOHKPYLEKDSQSEKLNLYLVYGTFTYFP 120
Db	58	EKYGIDVQFVEFNDYALPNEHVSQKGDIDANAQHKPYLDBDAKATLNLYLVYGTFTYFP 117
QY	121	LAGYSTKIKTIANELKDGATTIVPDPDSNLARALIIIEKGLIKLKDNNTLSTTIDYEN 180
Db	118	LAGYSRKIKKNNELQDGAKVVPNDPTNRGRALIIIEKGLIKLKDNANLLSTVIDYEN 177
QY	181	PKKLIVKEVDTSVAAARAIDVDLAVVNNNAVQGVGLTASENGVFVEDKSPVANIIVARA 240
Db	178	PKKLNITEVDTSVAAARALDVIDLAVVNNTVAGVGLNAQDNGVFVEDKSPVANIIVEST 237
QY	241	DNKDSKALQDFVKAYQTFDEVEAEAKKQFKQGVYIKGM 276
Db	238	DNKDSKAVQDFVKSYQTFEEVYQEGHGFQKGVYIKGM 273

```

RESULT 9
US-10-282-122A-67387
; Sequence 67387, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EILITA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20

```



```

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67387
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-67387
```

```

Query Match      64.2%; Score 888; DB 12; Length 276;
Best Local Similarity 60.2%; Pred. No. 1,1e-72;
Matches 168; Conservative 55; Mismatches 50; Indels 6; Gaps 3;

QY 1 MNEGKINGICALAGIALAGCSNQSNEPAIS---KTAQTIVGVMAEPQVAEVAQ 57
Db 1 MKRTKLFGLATVSAVALAGC-KDXPAAAPQGPAPARKLVGWTAEGVTEVAEK 58
QY 58 VAEEKNLVTEVEFNDYAMPNSAVSKGELDANAMQHKPELEKDSOEGKLNLIYGN 117
Db 59 IAEKKNIDIVKLEVEFEYQPNDAITKGDIDANAFQHKRYMDKEVQROGV-KIAIVGN 117
QY 118 VYELAGSTKIKTLNLEKDGATIAVNDPSNLARALILEKQGLIKLKNNTLFTTLDI 177
Db 118 VEFIAVSKKIKVSELQDQATVAVPNNPSNLGRALLLEKQGLIKLKPNSLFSISIV 177
QY 178 VENPKLVIEVDTSVAAPALDDVDLAVNNNTYAGVGLTASENGVEVEDKSPYVNIIV 237
Db 178 IENPKLVIEVDTSVAAPALDDVDLAVNNNTYAGVGLTASENGVEVEDKSPYVNIIV 237
QY 238 ARADNDSKALIDFVKAYQOTDEVEAEAKQFKDGVIRGW 276
Db 238 SREDNDSKALIDFVKAYQOTDEVEAEAKQFKDGVIRGW 276
```

```

RESULT 10
US-10-282-122A-76195
; Sequence 76195, Application US//10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
```

```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76195
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-76195
```

```

Query Match      57.4%; Score 793.5; DB 12; Length 271;
Best Local Similarity 56.5%; Pred. No. 4.7e-64;
Matches 156; Conservative 44; Mismatches 65; Indels 11; Gaps 3;

QY 3 FKGINGICALAGIALAGCSNQSNEPAISKTAQTIVGVMAEPQVAEVAQVAK 62
Db 5 FKTFAVGLALGSLAGCGGDEKDP-----NHRKGVIAEQAQVAEVAQVAK 56
QY 63 YNLTEVEFNDYAMPNSAVSKGELDANAMQHKPELEKDSOEGKLNLIYGNFTVYPLA 122
Db 57 YGLDVEITFNDYLLPENAISKGDIDANAFQHKRYLDOQIMDRGV-KLVSGKTFVYPLA 115
QY 123 GYSTKIKTLNLEKDGATIAVNDPSNLARALILEKQGLIKLKNNTLFTTLDI 182
Db 116 GYSKIKSLDELKQGSQVAVNPDPNTNIGRSLLIQKGLIKLKGVSILPTSLDIVENPK 175
QY 183 KLVIEVDTSVAAPALDDVDLAVNNNTYAGVGLTASENGVEVEDKSPYVNIIVARA 240
Db 176 NKLVELEAPQLPSSLDQAQALAVINTTASQGLIPAKDGIYVEDKSPYVNIIVTRE 235
QY 241 DNKDSKALIDFVKAYQOTDEVEAEAKQFKDGVIRGW 276
Db 236 DNKDAENVKFEVQAVQSDVEVEAANKVNGAVGW 271
```

```

RESULT 11
US-09-815-242-10050
; Sequence 10050, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: EUTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/257,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10050
; LENGTH: 271
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-815-242-10050
```

```

Query Match          56.6%; Score 782.5; DB 9; Length 271;
Best Local Similarity 55.1%; Pred. No. 4,7e-63;
Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
```

```

QY 3  FGTNGICALASGIALAGCSNQSNEPAAISKTAQTITKCYVMAQPEQVAEVAQVAKKK 62
    5  FKTFAAGALIGSLALVGGCGDEKDP-----NHIKGVIVGAEQOVAEVAQVAKKK 56
DB 63  YNLTFELVEFNDYAMPNSAVSKGELDANAMQHKRYLEKDSQEKGLNNLVIVGNTFVYPLA 122
    57  YGDLVELVTFNDYLPREHAIKSKGIDIDANAFQHKPYLDQQLKRGY-KLVAVGNTFVYPLA 115
QY 123 GYSTIKTLNELKQATITAVPNDPSNARALILLESQGLIKLKNNTLSTTLDIYENK 182
    116 GYSKKIKSLDELQSGQVAAPNDPTNLGRSLILLOKVGILKLDGVLPTVLDAVENPK 175
DB 183 KLVYKEVDTSVAARAIDD--VDLAVNNNYAGGVGLTASENGVEDEKSDSPYNIIVARA 240
    176 NLKIVLEAPQLPRLSDDAQIALAVINTTASQIGLTPACDGI FVEDEKSPYNIIVTRE 235
QY 241 DNKSKAIQDFVKAQYQDEVEAEAKKQFKQGVIRGW 276
    236 DNKDAENVKKEFQAYQSDVEVEAANKVFNGAVKGM 271
DB
```

```

RESULT 12
US-10-282-122A-43291
; Sequence 43291, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43291
; LENGTH: 271
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-282-122A-43291
```

```

Query Match          56.6%; Score 782.5; DB 12; Length 271;
Best Local Similarity 55.1%; Pred. No. 4,7e-63;
Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
```

```

QY 3  FGTNGICALASGIALAGCSNQSNEPAAISKTAQTITKCYVMAQPEQVAEVAQVAKKK 62
    5  FKTFAAGALIGSLALVGGCGDEKDP-----NHIKGVIVGAEQOVAEVAQVAKKK 56
DB 63  YNLTFELVEFNDYAMPNSAVSKGELDANAMQHKRYLEKDSQEKGLNNLVIVGNTFVYPLA 122
    57  YGDLVELVTFNDYLPREHAIKSKGIDIDANAFQHKPYLDQQLKRGY-KLVAVGNTFVYPLA 115
QY 123 GYSTIKTLNELKQATITAVPNDPSNARALILLESQGLIKLKNNTLSTTLDIYENK 182
    116 GYSKKIKSLDELQSGQVAAPNDPTNLGRSLILLOKVGILKLDGVLPTVLDAVENPK 175
DB 183 KLVYKEVDTSVAARAIDD--VDLAVNNNYAGGVGLTASENGVEDEKSDSPYNIIVARA 240
    176 NLKIVLEAPQLPRLSDDAQIALAVINTTASQIGLTPACDGI FVEDEKSPYNIIVTRE 235
QY 241 DNKSKAIQDFVKAQYQDEVEAEAKKQFKQGVIRGW 276
    236 DNKDAENVKKEFQAYQSDVEVEAANKVFNGAVKGM 271
DB
```

```

RESULT 13
US-10-282-122A-55826
; Sequence 55826, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

```

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55826
LENGTH: 271
TYPE: PR
ORGANISM: Enterobacter cloacae
US-10-282-122A-55826
```

```

Query Match 56.4%; Score 780.5; DB 12; Length 271;
Best Local Similarity 56.7%; Pred. No. 7.2e-63;
Matches 152; Conservative 43; Mismatches 62; Indels 11; Gaps 3;
```

```

QY 11 ALASGIALGCSNQSNEPAISKTAAQTITKVGMAPEQAVAEVAGVAKKNTLVEIV 70
DB 13 ALGSLALVGGCGDEKDP-----NHKVGIVGAEDQVAEVAQKAKKXGLDELIV 64
QY 71 EFNDAVMPNSAVSKEGELDANAMQHKPYLEKDSQEKLNLIYVGNTPFYVPLAGYSTKXT 130
DB 65 TENDYVLPNEALSKGIDANAFQHKPYLDQIQKDRGY-KLVVGNTPFYVPIAGYSKXTS 123
QY 131 LNLKGGATIAVNDPSNLARAILILEKQGLIKLKNNTLFTTLDIVENPKKLVKEVD 190
DB 124 LDELQGSQVAVNDPTNIGRSLLTLQKVGILKLDGVGLPTVLDVTENPKMLKIVELE 183
QY 191 TSAVAADID--VDLAVNNVAGVGLTASENGVEVEDKDSFYNNIYARADNDSKAI 248
DB 184 APQLPSLDDAQIALAVINTTYSQJGLTPAKDGIFFVEDKDSFYNNLIVREDNKAENV 243
QY 249 QDFVKAAYQTDVEVAEAKQKQFKDGVKGM 276
DB 244 KKFVQAYQSDDEVYQEAANKVFNGAVKGM 271
```

```

RESULT 14
US-10-282-122A-59600
Sequence 59600, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EUTRA.034A
CURRENT FILING DATE: 2003-02-20
```

```

PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 59600
LENGTH: 271
TYPE: PR
ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59600
```

```

Query Match 56.1%; Score 775.5; DB 12; Length 271;
Best Local Similarity 55.2%; Pred. No. 2.1e-62;
Matches 153; Conservative 45; Mismatches 68; Indels 11; Gaps 3;
```

```

QY 2 NFGKNGICALASGIALGCSNQSNEPAISKTAAQTITKVGMAPEQAVAEVAGVAKK 61
DB 4 NFKPAAAGALIGSLALVGGCGDEKDP-----NHKVGIVGAEDQVAEVAQKAKE 55
QY 62 KYNLTVELVEFNDAVMPNSAVSKEGELDANAMQHKPYLEKDSQEKLNLIYVGNTPFYVPL 121
DB 56 KYGLDELIVETPNDYVLPNEALSKGIDANAFQHKPYLDQIQKDRGY-KLVVGNTPFYVPI 114
QY 122 AGYSTKIKTLMELDGATIAVNDPSNLARAILILEKQGLIKLKNNTLFTTLDIVENP 181
DB 115 AGYSKIKSLDELPGSQIAVNDPTNIGRSLLTLQKVGILKLDGVGLPTSLDVENP 174
QY 182 KCLVKEVDTSVAARAID--VDLAVNNVAGVGLTASENGVEVEDKDSFYNNIYAR 239
DB 175 KNLKIVELEAPQLRSLDDAQIALAVINTTYSQJGLTPAKDGIFFVEDKDSFYNNLIVAR 234
QY 240 ADNDSKAIQDFVKAAYQTDVEVAEAKQKQFKDGVKGM 276
DB 235 EDNKAENVKKFVQAYQSDDEVYQEAANKVFNGAVKGM 271
```

```

RESULT 15
US-10-282-122A-68391
Sequence 68391, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EUTRA.034A
```

```
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ FILE REFERENCE: EUTRA.034A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 68391
/ LENGTH: 271
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
/ US-10-282-122A-68391

Query Match      54.8%; Score 757.5; DB 12; Length 271;
Best Local Similarity 52.7%; Pred. No. 9.1e-61;
Matches 147; Conservative 49; Mismatches 70; Indels 13; Gaps 4;

QY      1 MNFGKINGICALASGIALAGCSNOSNEPAISKTAQTIKYGMAGPEQVAEVAQVAK 60
DB      3 LKFKSLAVASALVGAALAGGGEKXP-----NHIRGVYISGEQVAEVAQVAK 54
QY      61 EKNVLTVELVEFNDYAMPNSAVSKGELDANAMQHKPELEKDSQEGKLN-MLVIVGNTFY 119
DB      55 DKYGLNVLVTFNDFVPMNESLSRGDIDINAFQHKPFL--DQIKDRNYKITAVGNTFY 112
QY      120 PLAGSTKIKTLNELKOGATIAVNDPSNLAARALLILEKQGLIKLXONTNLFSITLDIYE 179
DB      113 PLAGYSKKITPLADLPDAQVAIIPNDPTNLGRSILLILEKQGLIKLXONTNLFSITLDI 172
QY      180 NPKKLVKEVDTSVAARAID--VDLAVNNNNVAGVGLTASENGVEVEDKDSPYNITIV 237
DB      173 NPKNQLVELAPQPLPSLDQKITYLAVINTTASQVNLTPAKDGLFVEDKDSPYNITIV 232
QY      238 ARAADKSKAIQDFVKAQYQDVEAEAKKQFKQGVIXGM 276
DB      233 AREDKQDENVKKEFYQYQDDEVDAANKIFNGGAVXGM 271

RESULT 16
US-10-282-122A-78174
/ Sequence 78174, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
```

```
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: EUTRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 78174
/ LENGTH: 271
/ TYPE: PRT
/ ORGANISM: Yersinia pestis
/ US-10-282-122A-78174

Query Match      54.8%; Score 757.5; DB 12; Length 271;
Best Local Similarity 53.2%; Pred. No. 9.1e-61;
Matches 148; Conservative 48; Mismatches 71; Indels 11; Gaps 3;

QY      1 MNFGKINGICALASGIALAGCSNOSNEPAISKTAQTIKYGMAGPEQVAEVAQVAK 60
DB      3 LKFKSLAIAISALICTLTLVGCPTERKAP-----NHIKGVYISGEQVAEVAQVAK 54
QY      61 EKNVLTVELVEFNDYAMPNSAVSKGELDANAMQHKPELEKDSQEGKLNMLVIVGNTFY 120
DB      55 EKYGLDELVLTFNDFVPMNESLSRGDIDINAFQHKPFL--DQIKDRNYKITAVGNTFY 113
QY      121 LAGSTKIKTLNELKOGATIAVNDPSNLAARALLILEKQGLIKLXONTNLFSITLDIYE 180
DB      114 IAGYSKKITSLDELQPSQVALLPNDPTNLGRSILLILEKQGLIKLXONTNLFSITLDIYE 173
QY      181 PKKLVKEVDTSVAARAID--VDLAVNNNNVAGVGLTASENGVEVEDKDSPYNITIV 238
DB      174 PKNQLVELAPQPLPSLDQKITYLAVINTTASQVNLTPAKDGLFVEDKDSPYNITIV 233
QY      239 ARAADKSKAIQDFVKAQYQDVEAEAKKQFKQGVIXGM 276
DB      234 REDKQDENVKKEFYQYQDDEVDAANKIFNGGAVXGM 271

RESULT 17
US-10-282-122A-77174
/ Sequence 77174, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
```

```

APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO 77174
LENGTH: 275
TYPE: PRT
ORGANISM: Vibrio cholerae
US-10-282-122A-77174

```

```

Query Match          54.3%: Score 751.5; DB 12; Length 275;
Best Local Similarity 54.1%: Pred. No. 3.3e-60;
Matches 145; Conservative 51; Mismatches 63; Indels 9; Gaps 2;

```

```

QY 9 TCALSGIACGSSQNSNEPAISKTAAQTIVGVMAGPEQVAEVAQVAKKXNLTVE 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 17 IAAASTLILAGCGEAVD-----NNKXKIGVMAQEAVAEVAKVAKKXNLDVE 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 IVEFNDYAPNPAVSAKSGELDANAMQKPYLEKDSQKGLNLTIVGNTFFVYPLAGYSTKI 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 69 IVTFTDYTPNALDGDGSDANAFQHKPYLDKQIDRGY-KLAVGNFFVYPIAGYSKQI 127
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 KTLNELKQGTATVPPNPENLAPALITLEKQGLIKLKNLTLFSTLTIYENPKKLVYKE 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 KSVDELQDQARLAVPDPNLTLSRLILLOQOGLKLRDVLTLTVRDIVENPKKLEILB 187
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 189 VTSVAKAIDVDLAVVNNYAGOVGLTASNGVFEVDKDSPPYNIIVARADKDSKAI 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 188 LBAQCPRLSDVALLSTIITTYASSINLTPEKDGIFVKNSSPYNIIVAEAVNAENV 247
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 249 QDFVKAAYOTDEVEAEKAKQFQKGVIRGW 276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 248 QNFKAAYOTDEVEAKASEIFQGAIVKRW 275
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 18
US-09-815-242-11656
Sequence 11656, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

```

```

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11656
LENGTH: 240
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-815-242-11656

```

```

Query Match          53.8%: Score 743.5; DB 9; Length 240;
Best Local Similarity 60.0%: Pred. No. 1.4e-59;
Matches 144; Conservative 42; Mismatches 51; Indels 3; Gaps 2;

```

```

QY 39 IRVGMAGPEQVAEVAQVAKKXNLTVEVENDYAPNPAVSAKSGELDANAMQKPYL 98
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 IRVGVIAQEQVAEVAQVAKKXNLTVEVENDYAPNPAVSAKSGELDANAMQKPYL 61
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 99 EKDSQKGLNLTIVGNTFFVYPLAGYSTKI KTLNELKQGTATVPPNPENLAPALITLEK 158
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 62 QDQIDRGY-KLAVGNFFVYPIAGYSKQI KSLDELQPSGLIAPNDPNTLNGRLILLOQ 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 159 QGLIKLKNLTLFSTLTIYENPKKLVYKEDTVAARALTD--VDLAVNNYAGOVGL 216
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 VAILKLGQVGLTLPISLTIYENPKKLVLEAPQLPRSDDAQIALAVINTTYASQIGL 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 217 TASENGVFEVDKDSPPYNIIVARADKDSKAIODFVKAAYOTDEVEAEKAKQFQKGVIRGW 276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 TPANQGIPEGESESYVALIIVARENDKDAENVKRFVQAYQSDVEYEAANKIFNGAVKRW 240
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 19
US-10-282-122A-47704
Sequence 47704, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23

```

PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO: 47704
LENGTH: 295
TYPE: PRT
ORGANISM: Burkholderia cepacia
US-10-282-122A-47704

Query Match 40.6%; Score 562; DB 12; Length 295;
Best Local Similarity 44.5%; Pred. No. 7.1e-43;
Matches 118; Conservative 47; Mismatches 88; Indels 12; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVAQVAKKYNLTVEVE 71
DB 37 LAGALTLTSGAHAD-----KVIKVGTVAGDSVWQVQVAKKKEGKLVKYLE 88
QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEKGLNVLVGNTFVYPLAGYSTKIKTL 131
DB 89 FNYVQFNALDSDGDLANSFOHPYLDQVQKRGY-KIVSAGLYTSPIGVYSKKEPKSL 147
QY 132 NEKDGATIAVPNDPNTLALALILEKQGLIKLKDNNTL---FSTLTDIVENPKLVKE 188
DB 148 KDLPGKXKAVPDPENRRLILLOQGVITLTKAGAGTGNNAVLIAANPKKIKISE 207
QY 189 VDSVAAARADVDVLAVNNNTAGQVGLTASENGVFVEDKQSPYNTIIVAADNKDSKAI 248
DB 208 LDAQLEPRVLDVDAVINNTYALANLOPTKDALALBSLTSPYANLIIAVAKKDDOPW 267
QY 249 QDEVKAYQTDVEVEAEAKKQPKDGI 273
DB 268 KKLIVKAYQSPVEVEKFKQPKGSMV 292

RESULT 20

US-10-282-122A-51371
Sequence 51371, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO: 51371
LENGTH: 262
TYPE: PRT
ORGANISM: Bordetella pertussis
US-10-282-122A-51371

Query Match 39.8%; Score 551; DB 12; Length 262;
Best Local Similarity 43.1%; Pred. No. 6.1e-42;
Matches 118; Conservative 52; Mismatches 86; Indels 18; Gaps 5;

QY 1 MNRGKINGICALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVA-GOVA 59
DB 3 MNPVR---SALLASFLAGAG-----AKLVGATQVPHALILEVVPALAA 48
QY 60 KERNLTVELVEPNDYAMPNSAVSKGELDANAMQHKRYLEKDSOEKGLNVLVGNTFVY 119
DB 49 KEGVLDLKV--FTDYVQFNQLADKQIDANFFQHPYLDPTFNKDR-KTNLVSQGVHVE 105
QY 120 PLAGYSTKIKTLNELKQCATIAVPNDPNTLALALILEKQGLIKLKDNNTLFTSLTDIVE 179
DB 106 PEGYSKIKSLAEKQDATTIAPNDSNGRALLLOKQGLKXKADPSIIVATPIDIE 165
QY 180 NPKVLIVKEDVTSVAAARADVDVLAVNNNTAGQVGLTASENGVFVEDKQSPYNTIIVAR 239
DB 166 NPKVLKFELEAAMLPSEFDDLDLALINTYALAEAGLVPTPRDALFLEGADSPYANLVAAR 225
QY 240 ADNKSIAIDPFKAYQTDVEVEAEAKKQPKDGI 273
DB 226 PDNKDAPVAKLVNHLHSEAVRKTIKRYKGAIV 259

RESULT 21

US-10-282-122A-50375
Sequence 50375, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20

```
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 50375
/ LENGTH: 272
/ TYPE: PRT
/ ORGANISM: Burkholderia mallei
/ US-10-282-122A-50375
```

Query Match 39.7%; Score 549.5; DB 12; Length 272;

Best Local Similarity 42.5%; Pred. No. 8,8e-42;

Matches 114; Conservative 52; Mismatches 91; Indels 11; Gaps 3;

```
QY 9 ICALASGIALAGCSNOSNEPAAISKTAAQTIKVGWAGPEQAVAEVAGVAKKXNLTVE 68
DB 10 VSAVAGAAALSVS-----VGAQADPKVIVAGVDAQVQVQVAKKXGIDVK 62
QY 69 LVEFNDYAMPNSAVSGELDANAMQKPYLEDSQSEKGINLVYGVTFVYPLAGYSKI 128
DB 63 VIEFNDYVQPNALDSGDIDANSFQHPYLSQVQKRGV-KIVSAGLTVYISPIGVYSKF 121
QY 129 KTLMEKDGATIAVPNDPSNLARALILEKQGLIKLKONTNL---FSTLTDIVENPKLV 185
DB 122 KSLKELPAGAKYALPNDPENRERALLILQGVITLKAGAGCGSNALVDAENPKKXK 181
QY 186 IREVTISVAPARADVDYDLAVNNNNYAGVGLTASENGVFVEKDSFYNNIIVARAUNKDS 245
DB 182 LVELDAAGCPRYLISVDVAIVNTNYALANLOPTKDALSLSLSPFYNMLIAVAKKDQ 241
QY 246 KAIQDFKAYQTDVEVAEAKKQFKDGI 273
DB 242 PWTYKLVKAYQSEVYKFFITKQKGMV 269
```

RESULT 22

```
US-10-282-122A-61127
/ Sequence 61127; Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
```

```
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 61127
/ LENGTH: 256
/ TYPE: PRT
/ ORGANISM: Legionella pneumophila
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (227)..(227)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (241)..(241)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (244)..(244)
/ OTHER INFORMATION: X=any amino acid
/ NAME/KEY: MISC FEATURE
/ LOCATION: (248)..(248)
/ OTHER INFORMATION: X=any amino acid
/ US-10-282-122A-61127
```

Query Match 39.4%; Score 545; DB 12; Length 256;

Best Local Similarity 45.7%; Pred. No. 2,1e-41;

Matches 118; Conservative 37; Mismatches 91; Indels 12; Gaps 3;

```
QY 16 ILAGCSNOSNEPAAISKTAAQTIKVGWAGPEQAVAEVAGVAKKXNLTVEFNDY 75
DB 11 ISLVACS-----SKPAPNTLVIGTIAQPEHLEITFAQVAKKXNLTVEFNDY 61
QY 76 AMPNSAVSGELDANAMQKPYLEDSQSEKGINLVYGVTFVYPLAGYSKXITLNEK 135
DB 62 NLPNEALQDGSIDAVVYHLPYLKAAIISHGV-DLQAGRTFVYPMGIYSKKYTLSELP 120
QY 136 DGATIAVPNDPSNLARALILEKQGLIKLKONTNLFTLTDIVENPKLVKEVDTSVAA 195
DB 121 ENGIIAVPNDPENRERALLILEKKAHLITLKTNN--SGIQIBSNPKQFKKEDDAQLP 178
QY 196 RAIDVDYDLAVNNNNYAGVGLTASENGVFVEKDSFYNNIIVARAUNKDSAIQDFKAY 255
DB 179 RVLPDVDAIVNTTALPAGLSFSPSKDALFTGKDSFYNNIIVIRBDTEXRPQLELRYKAL 238
QY 256 QTDVEVAEAKKQFKDGI 273
DB 239 NSXVYKXKXNLFGEAI 256
```

RESULT 23

RESULT 29
US-09-815-242-5146
Sequence 5146, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA, 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5146
LENGTH: 259
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5146

Query Match 36.0%; Score 497.5; DB 9; Length 259;
Best Local Similarity 41.4%; Pred. No. 4.5e-37;
Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;
QY 9 ICLASGIALAGCSNQSNEPAISKTAAQTIKVGWAGPEQAVAE-VAGOVAKKKNLTV 67
Db 7 LTLAASV-----SVSASAEEKLIVATPIPHAEILILKFTLAKGCVLQI 53
QY 68 ELVEFNDAVMPNSAVSKGSLDANAMQHKPYLEKDSQEGKLNVLIVGNTFYYPPLAGYSTK 127
Db 54 KV--FTDYQVPVQVAKERLDANYFQTLPYLENFNKGGTGLTVVG-VHVEPFGYSRK 110
QY 128 IKTINELKDGATIAVNDPSNLARALILKEKGLIKLKDNTNLFSTTLDIVENPKKVIK 187
Db 111 YKSLALPBGATVATINBGSNSGRALLLQKAGLKLKDPNNALATPRDIAENFKNLKFK 170
QY 188 EVDTSVAARAIDVDVLAVNNNYAGOVGLTASENGVEVEDKDSPPYVNIIVAPADNKSKA 247
Db 171 ELKSALLPRVLDQVLDLINTVYALAEAKLNPAKDALVLEDRDSPPYVNIIVAPADNKSDA 230
QY 248 IODPVKAYOTDEVEAEAKKQKFDGVI 273
Db 231 LKLSAALTSPVKAIFIEKKYAGAVV 256

RESULT 30
US-10-282-122A-43469
Sequence 43469, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA, 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 78614
REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PAM.
SOFTWARE: Patent in version 3.1
SEQ ID NO 43469
LENGTH: 259
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-43469

Query Match 36.0%; Score 497.5; DB 12; Length 259;
Best Local Similarity 41.4%; Pred. No. 4.5e-37;
Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;
QY 9 ICLASGIALAGCSNQSNEPAISKTAAQTIKVGWAGPEQAVAE-VAGOVAKKKNLTV 67
Db 7 LTLAASV-----SVSASAEEKLIVATPIPHAEILILKFTLAKGCVLQI 53
QY 68 ELVEFNDAVMPNSAVSKGSLDANAMQHKPYLEKDSQEGKLNVLIVGNTFYYPPLAGYSTK 127
Db 54 KV--FTDYQVPVQVAKERLDANYFQTLPYLENFNKGGTGLTVVG-VHVEPFGYSRK 110
QY 128 IKTINELKDGATIAVNDPSNLARALILKEKGLIKLKDNTNLFSTTLDIVENPKKVIK 187
Db 111 YKSLALPBGATVATINBGSNSGRALLLQKAGLKLKDPNNALATPRDIAENFKNLKFK 170
QY 188 EVDTSVAARAIDVDVLAVNNNYAGOVGLTASENGVEVEDKDSPPYVNIIVAPADNKSKA 247
Db 171 ELKSALLPRVLDQVLDLINTVYALAEAKLNPAKDALVLEDRDSPPYVNIIVAPADNKSDA 230
QY 248 IODPVKAYOTDEVEAEAKKQKFDGVI 273
Db 231 LKLSAALTSPVKAIFIEKKYAGAVV 256

RESULT 31
US-10-282-122A-52976
Sequence 52976, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52976
; LENGTH: 263
; ORGANISM: Clostridium difficile
; US-10-282-122A-52976

Query Match      35.7%; Score 494; DB 12; Length 263;
Best Local Similarity 41.7%; Pred. No. 9.7e-37;
Matches 111; Conservative 50; Mismatches 87; Indels 18; Gaps 6;

QY 11 ALASGIALA--GCSNCSNEPAISKTAQTIKVGWAGPEQVAEVAQVAKER-YNLTV 67
DB 10 ALVASAIASVAGCSNKED-----KKILVGAASNPHAKILVAKPLKEKEGYDLEV 59
QY 68 ELVEENDYAMPNSAVKSGELDANAMQHKPYLEKDSQKGLNULVYGNTPFYVPLAGYSTK 127
DB 60 KI--FDYVJPNPALDEGSLDANFQHIPLLEBYEKGY-KLTYTSKVHIEPMGFISEK 116
QY 128 IKTLNELKDGATTAVPNDPSNLARALLILEKQGLIKLKNNTLFSFTTLDIVENPKLVYK 187
DB 117 VKSIDELKDGAVIAVPNDATNGARALKLAKNKLIEVKDGLI--TKXDITGNPNIVIK 174
QY 188 EVUDSVAAKAIIDVDVLAVVNNVNAAGVGLTASENGVFVEDKSPYVNIIVARADKDSKA 247
DB 175 EWMNEQLPTVLKVDGAVLNSNIALTANLNPTKDAIVIESSPSYVNIILACRENNKDSK 234
QY 248 IODFVAXYQTDVEAEAKQKQFQDGVIT 273
DB 235 IKALSEAMNSKEVKKFIQDEYKGSIV 260

```

RESULT 32
 US-10-282-122A-67627
 ; Sequence 67627, Application US/10282122A

```

; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67627
; LENGTH: 256
; ORGANISM: Pseudomonas putida
; US-10-282-122A-67627

Query Match      35.3%; Score 488; DB 12; Length 256;
Best Local Similarity 42.3%; Pred. No. 3.3e-36;
Matches 105; Conservative 48; Mismatches 89; Indels 6; Gaps 3;

QY 29 AAIKSTAAQTIKVGWAGP---EQAVAEVAQVAKERKNTLTVELVEFNDYAMPNSAVSKG 85
DB 9 AAVAAFSQAQADLTVAATPVPFAELINEFKQQLAKEGVLEKVK--EFTDYIQPNVQVAK 66
QY 86 ELDANAMQHKPYLEKDSQKGLNULVYGNTPFYVPLAGYSTKTIKTLNELKDGATTAVPND 145
DB 67 RLDANFPQHQPYLDENFAKQ--TSLVASVAGVHIEPLGVITIKIKLDELSGATVIEND 125
QY 146 PNLARALLILEKQGLIKLKNNTLFSFTTLDIVENPKLVYKEVDSVAAKAIIDVDVLAV 205
DB 126 ANNGRALLILDDKAGVILKDKNKNILSTVKQVAENPNKVNFRLEAATTIRAVTQVDAAL 185
QY 206 NNNNAAGVGLTASENGVFVEDKSPYVNIIVARADKDSKALQDPFKAYQTDVEAEAK 265
DB 186 INTNVALEKLANPEKDALIAGSDSPYVNIIVARADKDSDDKKAALAHSEFVKQFIT 245
QY 266 KQFKDGVIT 273
DB 246 EKYKGAIV 253

```

RESULT 33

```

1 RESULT 34
2 US-10-282-122A-57209
3 Sequence:57209, Application US/10282122A
4 Publication No. US2004002912A1
5
6 GENERAL INFORMATION:
7
8 APPLICANT: Wang, Liangsu
9 APPLICANT: Zamudio, Carlos
10 APPLICANT: Malone, Cheryl
11 APPLICANT: Haselbeck, Robert
12 APPLICANT: Ohlsen, Karl
13 APPLICANT: Zyskind, Judith
14 APPLICANT: Wall, Daniel
15 APPLICANT: Trawick, John
16 APPLICANT: Carr, Grant
17 APPLICANT: Yamamoto, Robert
18 APPLICANT: Forsyth, R.
19
20 APPLICANT: Xu, H.
21
22 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
23
24 FILE REFERENCE: ELITRA.034A
25
26 CURRENT APPLICATION NUMBER: US/10/282,122A
27
28 CURRENT FILING DATE: 2003-02-20
29
30 PRIOR APPLICATION NUMBER: 60/191,078
31
32 PRIOR FILING DATE: 2000-03-21
33
34 PRIOR APPLICATION NUMBER: 60/206,848
35
36 PRIOR FILING DATE: 2000-05-23
37
38 PRIOR APPLICATION NUMBER: 60/207,727
39
40 PRIOR FILING DATE: 2000-05-26
41
42 PRIOR APPLICATION NUMBER: 60/230,335
43
44 PRIOR FILING DATE: 2000-09-06
45
46 PRIOR APPLICATION NUMBER: 60/230,347
47
48 PRIOR FILING DATE: 2000-09-09
49
50 PRIOR APPLICATION NUMBER: 60/242,578
51
52 PRIOR FILING DATE: 2000-10-23
53
54 PRIOR APPLICATION NUMBER: 60/253,625
55
56 PRIOR FILING DATE: 2000-11-27
57
58 PRIOR APPLICATION NUMBER: 60/257,931
59
60 PRIOR FILING DATE: 2000-12-42
61
62 PRIOR APPLICATION NUMBER: 60/267,636
63
64 PRIOR FILING DATE: 2001-02-09
65
66 PRIOR APPLICATION NUMBER: 60/269,308
67
68
69 Remaining Prior Application data removed - See File Wrapper or PALM.
70
71 NUMBER OF SEQ ID NOS: 76614
72
73 SOFTWARE: PatentIn version 3.1
74
75 SEQ ID NO 57209
76
77
78 TYPE: PRT
79

```

Query Match	34.8%	Score 481	DB 12	Length 277
Best Local Similarity	36.8%	Pred. No. 1.6e-35		
Matches 102	Conservative 62	Mismatches 101	Indels 12	Gaps 5
QY	3	FGKINGICAMASGIALACCSNQSNEPAIKETAQITKGVMMAGPQAAEVAAGVAAEK	62	
Db	4	FSKLIGLIGVLA-FTIACGAGS-----VDYTKETIKLVGVATKNDWESEVDRIRKK-	56	
QY	63	VNLTVELVEFNDYAAPNPAVSKGELDANMMOHKYLEKSGOEKGLNNLTVGNTFYPPLA	122	
Db	57	-NIDQGVPEFDYTGPNPALAEKIDIDNAQHGFLEDDNKKENG-ITKLVISIGNTANAPLG	114	
QY	123	GYSTKTKTINLEKDGATTAVPNDPSNAPALILIEKOGILKDKONTNLPSTLIDIVENPK	182	
Db	115	IYANKTKDITIKIDGGEIAIENDEPTNGRALLILTQGLIKVDPAKQOLPYSIDITENKR	174	
QY	183	KLVYIKVEVDPSVAPARIDVDLAVVNNNNYAGOVGLTASENGVFE--DDSPVNNIVAR	239	
Db	175	QLKTEILDATQYRRALQVDASVINSGMAYDAGTTPKDAIFLEPNEBEAKFYVNIIVAR	234	
QY	240	ADNKKDSKALQDFVAKAYQTDVEAEAKQFQDGVIAKGM	276	

Db 235 EEDQENKLYQKVEEYQOEETKVIATSKGANVPAM 271

RESULT 35
US-10-206-576-30
Sequence 30, Application US/10206576
Publication No. US20030017495A1
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-10-206-576-30

Query Match 34.8%; Score 481; DB 12; Length 277;
Best Local Similarity 36.8%; Pred. No. 1,6e-35;
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

Cy 3 FGTNGICALASGIALAGCSNOSNEPAISKTPAQTIKVGVAAGEQVAEVAQVAK 62
Db 4 FSKLIGLIGVLA-FTLAGCAGS-----VMDTETEVKLGAVGTQNDWESVKRLKKK 56
Cy YNLVELVEFNDYAMPNSAVSKGELDANMOHPYLEKSOEKGNNLVIYGNFVYLA 122
Db 57 -NIDQVETFTDTPQRAALAEKEIDLNAPFOHIDPMDYNKHEG-TKLVSIGNTNAPLG 114
Cy GYSTKIKITLNELDGATIAVNDPSNLARALILLEKQGLIKKDNNTNLFSTLDIVENPK 182
Db 115 IYANKLDIRIKDGEIAPNDPTNGRALILQTAGLIRKVPKQOLPTVSDITERKR 174
Cy KLVIVEVDTSVAARIDVDLAVVNNVAGVGLTPASENGVE--DKDSPYNNIYAR 239
Db 175 QLKTEIDATQTRALQDPDASIVNGMAVDAGIYPPDKAIFLEFVNEKAPYNNIYAR 234
Cy 240 ADNKDSKAIQDPFKAIQTDEVEAKQKQFYDGVYIGW 276
Db 235 EEDQENKLYQKVEEYQOEETKVIATSKGANVPAM 271

RESULT 36
US-10-282-122A-54418
Sequence 54418, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Olsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EPIRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 54418
LENGTH: 257
TYPE: PRT
ORGANISM: Campylobacter jejuni
US-10-282-122A-54418

Query Match 34.5%; Score 477.5; DB 12; Length 257;
Best Local Similarity 41.5%; Pred. No. 3e-35;
Matches 100; Conservative 50; Mismatches 78; Indels 13; Gaps 5;

Cy 30 AISKTAQTIKVGVAAGEQVAEVAQVAK-EKYNLVELVEFNDYAMPNSAVSKGEL 87
Db 15 SINAAALFTITVAATPVPH--AEILQVKNPDEKQGYKLEIKFTDYVLPNLAVNGEA 71
Cy DANMOHPYLEKSOEKGNNLVIYGNFVYPLAGSTIKITLNELDGATIAVNDPS 147
Db 72 DANFQHTPYLEEFNKNKG-TKLKVAIHIEPAVYSKYSGLDIDKGVKIALPNDPT 130
Cy MLARALILLEKQGLIKKDNNTNLFSTLDIVENPKQLVIEVDTSVAARIDVDLAVN 207
Db 131 NERRALDIIIAKGLVKEKDKA--LKTPLDIIDNPKIKFVELKPAOLPRALNDVDPAVIN 188
Cy NNVAQVGLTPASENGVEFEDSDSYNNIYARADNDSKAIQDPFKAYQOTDEVEAKRQ 267
Db 189 SNTALSNANPNKQSVFLDEKSERYANILVVRVGHENDPKIKALQALQSDKI-----KQ 243
Cy 268 F 268

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rapb

Page 19

Db 244 F 244

```
RESULT 37
US-10-282-122A-52496
; Sequence 52496, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52496
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-282-122A-52496

Query Match 34.2%; Score 472.5; DB 12; Length 270;
Best Local Similarity 40.4%; Pred. No. 9.2e-35;
Matches 105; Conservative 54; Mismatches 86; Indels 13; Gaps 6;

Cy 16 IALAGCSNOSNEPAISKTAQITKYGVVAGPEQVAEYAGQVAK--EKNLTVELVAFEN 73
Db 19 VALTGGGS--SKEAKERK---TIYVGATPEPH--AEILKKVKXELERKGYTLEIKFT 70
Cy 74 DYAMPASVSKGELDANAMQHKPYLEKDSOEKGLNNLVIGNTFVYPLAGYSTKIKTINE 133
Db 71 DYTTPPTALQDGDIDANFYOHITPYLEEFYKER-KTDLSITVYVHLEPBGVSVSKTKDOLKE 129
Cy 134 LKGGATITAVPNDPSNLARALILLEKQGLIKDKONTNLFSTLIDYENRKVITKEVDTSV 193
Db 130 LKNGATISIPSDTNSRALKLEKRGITIKKEGE--LWSKMDITKPKNKIKIEBLDAAQ 187
Cy 194 AARALIDVDLAVVNNVAGQVGLTASENGVEVEDKDSFYVNIIVARANDKSKALIDPFVK 253
Db 188 LPTLTDVDAVAATNTNYVAVPANNPLPKDMLALIESKDSFYAVNIIVKTENKAAEYTKALDE 247
```

Cy 254 AYOTDEVEAEAKKORFDGVI 273
Db 248 AINSEIKYIEQYKGAITL 267

```
RESULT 38
US-10-282-122A-52838
; Sequence 52838, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52838
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-282-122A-52838
```

```
Query Match 34.1%; Score 472; DB 12; Length 270;
Best Local Similarity 40.9%; Pred. No. 1e-34;
Matches 112; Conservative 46; Mismatches 108; Indels 8; Gaps 5;

Cy 1 MNFGKINGI-CALASGIALAGCSNOSNEPAISKTAQITKYGVVAGPEQVAEYAGQVA 59
Db 1 MKKKRLTGLTSLVFTLIGVGCSSKESKRYVNDK---KTIYVGATPPVAGELIKVA-QPL 56
Cy 60 KKKVNLTVLAVFENYVAMNSAVSKGELDANAMQHKPYLEKDSOEKGLNNLVIGNTFVY 119
Db 57 LKKKGYLBIKFTTYVTPNTLRNREIDANFYOHITPYLEEFYKER-KTELEAVKIKYIS 115
Cy 120 PLAGYSTKIKTINEKQATTAVPNDPSNLARALILLEKQGLIKDKONTNLFSTLIDIVE 179
Db 116 PLATYSNKTKLEGLKDGATTAVPNDPTNERPALRLLEKSGILIKKGGDTL--TKGDIAE 173
Cy 180 NPKLVITKEVDTSVAABAIDVDLAVVNNVAGQVGLTASENGVEVEDKDSFYVNIIVAR 239
```

Db 174 NKKNIKKEIDAPQLPILNDVAVINTNYALAEATLNELKOSILEDSDSYANVIANK 233
QY 240 ADNRKSKAIODFKAYOTDEVEAEAKKOFKQDGI 273
Db 234 KEDKXAKIRALSEALTSKEVKOFINEKXGVVI 267

RESULT 39

US-10-282-122A-54417
Sequence 54417, Application US/10282122A
Publication No. US20040029128A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO 54417
LENGTH: 256
TYPE: PRT
ORGANISM: Campylobacter jejuni
US-10-282-122A-54417

Query Match 33.8%; Score 468; DB 12; Length 256;
Best Local Similarity 42.2%; Pred. No. 2,2e-34;
Matches 103; Conservative 52; Mismatches 83; Indels 6; Gaps 4;
QY 31 ISKTAQTIKYGVAEQAQVAEVAQVAKK-KYLVTELVFENDYAMPNSAVSKGELDA 89
Db 15 LSSFAQNTITGATPMPFSGILMKDPEKNGYEL--KIVFESDYILPNPALEKEKEDA 72
QY 90 NAMOHKRYLEKDSOEGLNNLVVGNTPFYVPLAGYSTKIKTLNELKQDQATIAVPDPBNL 149
Db 73 NLIOHKEFEEYVWKKG-SVLIATTTVPLIAPVGVYSKIKLNLKENTLKGARVAIPVDATNE 131
QY 150 ARAILILEKQGLIKLNDNTNLFSTLIDIVENPKULVIKEVDTSVARAIDVDLAIVANN 209
Db 132 SRALELEKAKLIEL--NKNTLTKPHIDINKPKKLFIEKKAQQLPALDVIDAIINSN 189

QY 210 YAGOVGLTASNGVFEVDESPYVNIIVARADNKSKAIDOFKAYOTDEVEAEAKKOFK 269
Db 190 FALGALNSKOTIFREDKNSPYVYVVESEGNSEKTKYIDELIASDKKALINHHX 249
QY 270 DGYI 273
Db 250 DILI 253

RESULT 40

US-09-815-242-11460
Sequence 11460, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11460
LENGTH: 271
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-815-242-11460

Query Match 33.8%; Score 467.5; DB 9; Length 271;
Best Local Similarity 42.1%; Pred. No. 2.7e-34;
Matches 99; Conservative 46; Mismatches 85; Indels 5; Gaps 3;
QY 39 IYGVWAGEQAQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAYOKRYL 98
Db 39 LRYGANPYPHAQILOSVDLDEK-KYLVIVSFYDVLNMLANDSLDANFYFORHPYL 97
QY 99 EKDSOEKGLNNLVVGNTPFYVPLAGYSTKIKTLNELKQDQATIAVPDPBNL 158
Db 98 DFNIDRKM-KHVGANIHVEPLRFYSKTIIDINKLKGSAIVAPNDPAOGGALLILHK 156
QY 159 QGLIKLNDNTNLFSTLIDIVENPKULVIKEVDTSVARAIDVDLAIVANNYAGQGLTA 218
Db 157 QGLIKLNDNTNLFSTLIDIVENPKULVIKEVDTSVARAIDVDLAIVANNYAGQGLTA 216
QY 219 SENGAVFEDKQSPYVNIIVARADNKSKAIDOFKAYOTDEVEAEAKKOFKQDGI 273
Db 217 A--LFSKDSFPYANLNAEDNAQDEALITLIEALQSEKTRKFLIDTYKGAII 268

Search completed: June 16, 2004, 11:14:11
Job time: 51 secs

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rapb

Page 21



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:00:16 ; Search time 60 seconds
(without alignments)
1259.719 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 1383
Sequence: 1 MNFKINGICALASGIALAG.....TDEVEAAKQKQGVKXGW 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_293an04.*
1: geneseqp19806.*
2: geneseqp19906.*
3: geneseqp20008.*
4: geneseqp20018.*
5: geneseqp20028.*
6: geneseqp20038.*
7: geneseqp20038.*
8: geneseqp20046.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1383	100.0	276	4	ABU20104 Moraxella
2	1368	98.9	276	6	AAO17582 Moraxella
3	1368	98.9	276	6	ABU35462 Moraxella
4	1007.5	72.8	273	5	ABU76073 Haemophil
5	1006.5	72.8	273	5	AAU35474 Haemophil
6	1006.5	72.8	273	5	ABU76074 Haemophil
7	1006.5	72.8	273	5	ABU30316 Moraxella
8	1004	72.6	272	5	AAO17813 Moraxella
9	888	64.2	276	6	ABU39463 Moraxella
10	793.5	57.4	271	6	ABU48271 Moraxella
11	782.5	56.6	271	6	AAU34457 Moraxella
12	782.5	56.6	271	6	ABU15367 Moraxella
13	780.5	56.4	271	6	ABU27902 Moraxella
14	775.5	56.1	271	6	ABU31676 Moraxella
15	767.5	55.5	272	6	ABU67576 Moraxella
16	757.5	54.8	271	6	ABU50250 Moraxella
17	757.5	54.8	271	6	ABU40467 Moraxella
18	751.5	54.3	275	6	ABU49280 Moraxella
19	743.5	53.8	240	4	AAU36063 Klebsiell
20	562	40.6	295	6	ABU19780 Moraxella
21	551	39.8	262	6	ABU23447 Moraxella
22	549.5	39.4	272	6	ABU22451 Moraxella
23	545	39.4	256	6	ABU33293 Moraxella
24	540	39.0	268	6	ABU21332 Moraxella
25	527.5	38.1	270	6	ABU21415 Moraxella

26	526.5	38.1	265	6	ABU23277 Moraxella
27	509.5	36.8	286	5	ABU53625 Moraxella
28	506.5	36.6	261	6	ABU3768 Moraxella
29	502.5	36.3	260	4	ABU36517 Moraxella
30	502.5	36.3	260	6	ABU38920 Moraxella
31	497.5	36.0	259	4	AAU33650 Moraxella
32	497.5	36.0	259	6	ABU15545 Moraxella
33	494	35.7	253	6	ABU25052 Moraxella
34	488	35.3	256	6	ABU39703 Moraxella
35	487	35.2	257	6	ABU41435 Moraxella
36	481	34.8	277	2	AAU00024 Moraxella
37	481	34.8	277	5	ABU43243 Moraxella
38	481	34.8	277	6	ABU88271 Moraxella
39	481	34.8	277	6	ABU29285 Moraxella
40	481	34.8	277	6	ABU13522 Moraxella
41	477.5	34.5	257	6	ABU26494 Moraxella
42	472.5	34.2	270	6	ABU24572 Moraxella
43	472	34.1	256	6	ABU24914 Moraxella
44	468	33.8	256	6	ABU26493 Moraxella
45	467.5	33.8	271	2	AAU90023 Moraxella

ALIGNMENTS

RESULT 1
ID AAB20104 standard; protein; 276 AA.
XX
AC AAB20104;
XX
DT 23-APR-2001 (first entry)
XX
DE Moraxella catarrhalis BAS11 protein.
XX
KW BAS11; infection; otitis media; pneumonia; diagnosis; therapy;
KW antibacterial; antimicrobial; vaccine.
XX
OS Moraxella catarrhalis.
XX
PN WC2001.00637-A1.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-EP005552.
XX
PR 25-JUN-1999; 99GB-00014945.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Thonard J;
XX
DR WPT; 2001-123013/13.
DR N-PSDB; AAF30040.
XX
PT New BAS11 polypeptides of Moraxella catarrhalis useful for diagnostic,
PT prophylactic and therapeutic purposes against microbial diseases,
PT preferably bacterial infections.
PS
PS Claim 1; Page 63; 79pp; English.
CC The present sequence is that of BAS11 protein from Moraxella
CC catarrhalis strain MC2931 (ATCC 43617), a causative agent of otitis media
CC in children and pneumonia in adults. The invention provides BAS11
CC polypeptides, and polynucleotides encoding them, as well as expression
CC vectors, host cells and methods for producing BAS11 polypeptides using a
CC recombinant method. Also claimed is a vaccine composition comprising a
CC BAS11 polypeptide, an immunogenic fragment of a BAS11 polypeptide, or
CC a polypeptide having at least 85% amino acid sequence identity to
CC BAS11, or comprising a polynucleotide encoding such a polypeptide. A
CC claimed method of diagnosing a Moraxella infection involves identifying a
CC BAS11 polypeptide or antibody. A claimed therapeutic composition useful
CC in treating humans with M. catarrhalis infection comprises at least 1

CC antibody directed against a BASH11 polypeptide. BASH11 polypeptides
CC also have utility in raising specific antibodies, and in screening for
CC antibacterial drugs
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1363; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.4e-119; Mismatches 0; Gaps 0;
Matches 276; Conservative 0; Indels 0;

QY 1 MNFGKINGICALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEQAVAEVAGQVAK 60
DB 1 MNFGKINGICALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEQAVAEVAGQVAK 60
QY 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
DB 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
QY 121 LAGSTKIKTLNELKOGATIAVPNDPSNLARALILEKQGLIKLNDNTLFFSTLIDIVEN 180
DB 121 LAGSTKIKTLNELKOGATIAVPNDPSNLARALILEKQGLIKLNDNTLFFSTLIDIVEN 180
QY 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKQSPYVNIIVARA 240
DB 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKQSPYVNIIVARA 240
QY 241 DNKSKAIDQFVAKYQTDVEAEAKKQFKDGVYIKGM 276
DB 241 DNKSKAIDQFVAKYQTDVEAEAKKQFKDGVYIKGM 276

RESULT 2
AA017582
ID AA017582 standard; protein; 276 AA.

XX AA017582;
XX 19-JUL-2002 (first entry)
XX
DE M catarrhalis MCA101896 protein SEQ ID NO: 44.
XX
KM Moraxella; vaccine; respiratory tract infection; anti-inflammatory;
XX auditory; antibacterial; otitis media; sinusitis; pneumonia.
XX
OS Moraxella catarrhalis.
XX
PN WO200218595-A2.

N-2

XX PD 07-MAR-2002.
XX 28-AUG-2001; 2001WO-CA001221.
XX
XX 28-AUG-2000; 2000US-0228294P.
XX 28-AUG-2000; 2000US-0228295P.
XX 28-AUG-2000; 2000US-0228296P.
XX 28-AUG-2000; 2000US-0228438P.
XX 28-AUG-2000; 2000US-0228439P.
XX 28-AUG-2000; 2000US-0228440P.
XX 28-AUG-2000; 2000US-0228441P.
XX 28-AUG-2000; 2000US-0228442P.
XX 28-AUG-2000; 2000US-0228443P.
XX 28-AUG-2000; 2000US-0228511P.
XX 28-AUG-2000; 2000US-0228512P.
XX 28-AUG-2000; 2000US-0228742P.
XX 28-AUG-2000; 2000US-0228743P.
XX 28-AUG-2000; 2000US-0229465P.
XX 28-AUG-2000; 2000US-0229474P.
XX 28-AUG-2000; 2000US-0229475P.
XX 28-AUG-2000; 2000US-0229478P.
XX 28-AUG-2000; 2000US-0229479P.
XX 28-AUG-2000; 2000US-0229603P.
XX 28-AUG-2000; 2000US-0229604P.
XX 28-AUG-2000; 2000US-0229605P.

PR 05-SEP-2000; 2000US-0229806P.
PR 05-SEP-2000; 2000US-0229809P.
PR 05-SEP-2000; 2000US-0229811P.
PR 05-SEP-2000; 2000US-0230214P.
PR 05-SEP-2000; 2000US-0230250P.
PR 05-SEP-2000; 2000US-0230252P.
XX
XX (AVET) AVENTIS PASTEUR LTD.
XX
XX Loosmore S, Wang J, Bradley B, Ochs M, Yang Y;
XX WPI; 2002-404555/43.
XX DR N-PSDB; AAL46514.
XX

PT Moraxella polypeptide and polynucleotides useful as vaccine for
PT immunizing a host e.g. humans against disease e.g. otitis media,
PT pneumonia, caused by infection of the bacteria.

XX Claim 28; Fig 43; 277pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC proteins from Moraxella catarrhalis. These can be used to produce
CC vaccines which protect against M. catarrhalis infection, which can cause
CC otitis media, respiratory infection, sinusitis, and pneumonia. The
CC present sequence is a protein of the invention
XX

SQ Sequence 276 AA;

Query Match 98.9%; Score 1368; DB 5; Length 276;
Best Local Similarity 99.3%; Pred. No. 3.5e-118; Mismatches 2; Indels 0; Gaps 0;
Matches 274; Conservative 0; Mismatches 2;

QY 1 MNFGKINGICALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEQAVAEVAGQVAK 60
DB 1 MNFGKINGICALASGIALAGCSNQSNEPAISKTAQTIKVGWAGPEQAVAEVAGQVAK 60
QY 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
DB 61 EKNLTVLVEFNDYMPNSAVSKGELDANAMQHKRYLEKDSQEGKLNLLVVGNTFYVP 120
QY 121 LAGSTKIKTLNELKOGATIAVPNDPSNLARALILEKQGLIKLNDNTLFFSTLIDIVEN 180
DB 121 LAGSTKIKTLNELKOGATIAVPNDPSNLARALILEKQGLIKLNDNTLFFSTLIDIVEN 180
QY 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKQSPYVNIIVARA 240
DB 181 PKKLVIKEVDTSVAARAIDVDLAVVNNNYAGGVGLTASENGVFVEDKQSPYVNIIVARA 240
QY 241 DNKSKAIDQFVAKYQTDVEAEAKKQFKDGVYIKGM 276
DB 241 DNKSKAIDQFVAKYQTDVEAEAKKQFKDGVYIKGM 276

RESULT 3

ABU35462
ID ABU35462 standard; protein; 276 AA.

XX ABU35462;
XX 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #20989.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Moraxella catarrhalis.
XX OS WO200277183-A2.
XX PN 03-OCT-2002.
XX PD 21-MAR-2002; 2002WO-US009107.
XX PF

21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohleen KL, Zyskind JW,
Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
WPI; 2003-029926/02.
N-PSDB; ACA39332.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 63386; 1766P; English.

The invention relates to an isolated nucleic acid comprising any one of
the 623 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 276 AA;

Query Match 98.9%; Score 1368; DB 6; Length 276;
Best Local Similarity 99.3%; Pred. No. 3.5e-118;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MNFGKINGICALAGSIALAGCSNOSNEPAISRTAAQITKVGWAGPEQVAEVAQYAK 60
1 MNFGKIFGICALAGSIALAGCSNOSNEPAISRTAAQITKVGWAGPEQVAEVAQYAK 60
1 EKNLTVELVERNDYAMPNSAVSKGSLDANAOHRYEKSDQEGLNLTVGNTFYYP 120
61 EKNLTVELVERNDYAMPNSAVSKGSLDANAOHRYEKSDQEGLNLTVGNTFYYP 120
121 LAGYSTKIKTINELKDGATIAVNDPSNLARALLILEKQGLIKLKDNTLFTSTLDIYEN 180
121 LAGYSTKIKTINELKDGATIAVNDPSNLARALLILEKQGLIKLKDNTLFTSTLDIYEN 180
181 PKXLVIKVDTSVAAPALDDVDLAVNNNYAGVGGLTASENGVFVEDKSPVNTIIVAA 240
181 PKXLVIKVDTSVAAPALDDVDLAVNNNYAGVGGLTASENGVFVEDKSPVNTIIVAA 240
181 PKXLVIKVDTSVAAPALDDVDLAVNNNYAGVGGLTASENGVFVEDKSPVNTIIVAA 240

241 DNKSKAIQDFVAKYOTDEVEAEAKKQFGKGVTKGW 276
241 DNKSKAIQDFVAKYOTDEVEAEAKKQFGKGVTKGW 276

RESULT 4
ABB76073
ID ABB76073 standard; protein; 273 AA.
AC ABB76073;
DT 15-JUL-2002 (first entry)
XX Haemophilus influenzae BASB202 protein.
XX BASB202; NTH; infection; vaccine; genetic immunisation; auditory;
XX antiinflammatory; antibacterial; immunostimulant; otitis media.
XX Haemophilus influenzae.

Key Location/Qualifiers

FT Region 1..15
FT /note= "potential T-helper cell epitope"
FT Region 21..28
FT /note= "potential B-cell epitope"
FT Region 34..42
FT /note= "potential T-helper cell epitope"
FT Region 43..47
FT /note= "potential B-cell epitope"
FT Region 56..60
FT /note= "potential B-cell epitope"
FT Region 66..74
FT /note= "potential T-helper cell epitope"
FT Region 69..74
FT /note= "potential B-cell epitope"
FT Region 78..84
FT /note= "potential B-cell epitope"
FT Region 89..97
FT /note= "potential T-helper cell epitope"
FT Region 93..104
FT /note= "potential B-cell epitope"
FT Region 107..126
FT /note= "potential T-helper cell epitope"
FT Region 121..127
FT /note= "potential B-cell epitope"
FT Region 129..134
FT /note= "potential B-cell epitope"
FT Region 141..148
FT /note= "potential B-cell epitope"
FT Region 150..169
FT /note= "potential T-helper cell epitope"
FT Region 176..182
FT /note= "potential B-cell epitope"
FT Region 181..189
FT /note= "potential T-helper cell epitope"
FT Region 200..208
FT /note= "potential T-helper cell epitope"
FT Region 224..228
FT /note= "potential B-cell epitope"
FT Region 229..240
FT /note= "potential T-helper cell epitope"
FT Region 236..244
FT /note= "potential B-cell epitope"
FT Region 249..256
FT /note= "potential B-cell epitope"
FT Region 259..267
FT /note= "potential B-cell epitope"

WO200224729-A2.

28-MAR-2002.

PF 18-SEP-2001; 2001WO-EP010979.
 XX
 XX 19-SEP-2000; 2000GB-00022992.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Thonard J;
 XX
 XX WPI, 2002-383180/41.
 DR N-PSDB; ABL56953, ABL56954, ABL56956.
 XX
 PT New isolated BASB202 polypeptide of nontypeable Haemophilus influenzae
 PT useful for diagnosing a disease and in generating an immune response in
 PT an animal.
 PS Claim 3; Page 75; 90pp; English.
 XX
 XX The present sequence is the protein sequence for the BASB202 protein of
 CC nontypeable Haemophilus influenzae (NTHi) strain 3224A (ATCC PT-1816)
 CC isolated from an otitis media patient in the USA. NTHi strain 3219C
 CC isolated from an otitis media patient in the USA. NTHi strain 810956
 CC isolated from a meningitis patient in the Netherlands, and NTHi strain
 CC A940184, a carrier strain from the Netherlands. Variability analysis of
 CC BASB202 protein among NTHi strains revealed 99-100% sequence identity.
 CC The invention provides BASB202 polypeptides and polynucleotides, vectors,
 CC host cells, and methods for producing the polypeptides by recombinant
 CC methods. Claimed vaccine compositions comprise a BASB202 polypeptide or
 CC polynucleotide. A claimed method of diagnosing NTHi infection involves
 CC identifying a BASB202 polypeptide or an antibody that is immunospecific
 CC for the polypeptide. A claimed therapeutic composition useful for
 CC treating humans with NTHi disease comprises an antibody directed against
 CC a BASB202 polypeptide. NTHi diseases include otitis media, pneumonia,
 CC sinusitis, nosocomial infections, invasive disease, chronic otitis media
 CC with hearing loss, fluid accumulation in the middle ear, auditory nerve
 CC damage, delayed speech learning, infection of the upper respiratory
 CC tract, and inflammation of the middle ear. B-cell epitopes and T-helper
 CC cell epitopes from BASB202 may also be useful in vaccine compositions
 XX
 XX Sequence 273 AA;
 SQ
 Query Match 72.8%; Score 1007.5; DB 5; Length 273;
 Best Local Similarity 71.7%; Pred. No. 8.7e-85;
 Matches 198; Conservative 27; Mismatches 48; Indels 3; Gaps 1;
 QY 1 MNEFKINGICALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVAQVAK 60
 DB 1 MKUKQFAITATASALVLTGCKEDKKPEAA--AAPLKIKVGMGSEPHQVAEIAKVAK 57
 QY 61 EKNLTVLVEFNDYAMPNSAVSKGELDANAMQHKRYLEKDSQEKLNLLVYGNTPYFP 120
 DB 58 EKTGLDVQFVEFNDYALPNEAVSKGDLNANMOKRXYLDEDAKAKNLNLLVYGNTPYFP 117
 QY 121 LAGYSTIKITNELKDGATIAVNDPSNLARALILIEKQGLIKLKNNTLFTSLDIVEN 180
 DB 118 LAGYSKKIKVNNELQDGAQKVVNDPTNNGRALILIEKQGLIKLKNNTLFTSLDIVEN 177
 QY 181 PKULVKEVDTSVAARAIDVDVLA VNNYAGOVGLTASENGVFVEDKDSPTVNIIVARA 240
 DB 178 PKKLNTEVDTSVAARAIDVDVLA VNNYAGOVGLTASENGVFVEDKDSPTVNIIVARA 237
 QY 241 DNKDSKAIODFVAKYOTDEVEAEAKQFQGVYIKGM 276
 DB 238 DNKDSKAYODFVAKSYOTEEVQEAQKHFQDGVYIKGM 273
 RESULT 5
 AAU35474
 ID AAU35474 standard; protein; 273 AA.
 XX
 AC AAU35474;
 XX
 DT 14-FEB-2002 (first entry)
 XX

DE Haemophilus influenzae cellular proliferation protein #115.
 XX
 XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 XX Haemophilus influenzae.
 OS
 XX WO200170955-A2.
 EN
 XX 27-SEP-2001.
 PD
 XX 21-MAR-2001; 2001WO-US009180.
 XX
 PF 21-MAR-2000; 2000US-0191078P.
 XX 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Haseelbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 N-PSDB; AAS53333.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 PT
 XX Example 3; SEQ ID NO 11067; 511pp; English.
 PS
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 XX
 XX Sequence 273 AA;
 SQ
 Query Match 72.8%; Score 1006.5; DB 4; Length 273;
 Best Local Similarity 71.4%; Pred. No. 1.1e-84;
 Matches 197; Conservative 28; Mismatches 48; Indels 3; Gaps 1;
 QY 1 MNEFKINGICALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVAQVAK 60
 DB 1 MKUKQFAITATASALVLTGCKEDKKPEAA--AAPLKIKVGMGSEPHQVAEIAKVAK 57
 QY 61 EKNLTVLVEFNDYAMPNSAVSKGELDANAMQHKRYLEKDSQEKLNLLVYGNTPYFP 120
 DB 58 EKTGLDVQFVEFNDYALPNEAVSKGDLNANMOKRXYLDEDAKAKNLNLLVYGNTPYFP 117
 QY 121 LAGYSTIKITNELKDGATIAVNDPSNLARALILIEKQGLIKLKNNTLFTSLDIVEN 180
 DB 118 LAGYSKKIKVNNELQDGAQKVVNDPTNNGRALILIEKQGLIKLKNNTLFTSLDIVEN 177
 QY 181 PKULVKEVDTSVAARAIDVDVLA VNNYAGOVGLTASENGVFVEDKDSPTVNIIVARA 240

D6		178	PXKLTINTEVDTSVAALADDDVLAVNNNTYAQGVGLANNQDGSVFEDKDSPYNIISRT	23
Oy		241	DNKDSKAIDPFVAKAYOTDEVEAEAKKFQFDGCIKW	276
			: : :	
D6		238	DNKDSKAVDPFIKSXYOTEVEYGEAQKHFPDGIVKGM	273
RESULT 6				
ID	ABB76074	standard; protein; 273 AA.		
XX	AC	ABB76074;		
XX	AB	ABB76074;		
DT	15-JUL-2002	(first entry)		
XX				
DE		Haemophilus influenzae BABS202 protein.		
XX				
KW	BABS202; NTH;	infection; vaccine; genetic immunisation; auditory;		
KM	antihistaminatory; antibacterial;	immunostimulant; otitis media.		
XX				
OS	Haemophilus influenzae.			
Key		Location/Qualifiers		
FT	Region	1..15	/note= "potential T-helper cell epitope"	
FT	Region	21..28	/note= "potential B-cell epitope"	
FT	Region	34..42	/note= "potential T-helper cell epitope"	
FT	Region	43..47	/note= "potential B-cell epitope"	
FT	Region	56..60	/note= "potential B-cell epitope"	
FT	Region	66..74	/note= "potential T-helper cell epitope"	
FT	Region	69..74	/note= "potential B-cell epitope"	
FT	Region	78..84	/note= "potential B-cell epitope"	
FT	Region	89..97	/note= "potential T-helper cell epitope"	
FT	Region	93..104	/note= "potential B-cell epitope"	
FT	Region	107..126	/note= "potential T-helper cell epitope"	
FT	Region	121..127	/note= "potential B-cell epitope"	
FT	Region	128..134	/note= "potential B-cell epitope"	
FT	Region	141..148	/note= "potential B-cell epitope"	
FT	Region	150..169	/note= "potential T-helper cell epitope"	
FT	Region	176..182	/note= "potential B-cell epitope"	
FT	Region	181..189	/note= "potential T-helper cell epitope"	
FT	Region	200..208	/note= "potential T-helper cell epitope"	
FT	Region	224..228	/note= "potential B-cell epitope"	
FT	Region	229..240	/note= "potential T-helper cell epitope"	
FT	Region	236..244	/note= "potential B-cell epitope"	
FT	Region	249..256	/note= "potential B-cell epitope"	
FT	Region	259..267	/note= "potential B-cell epitope"	
XX				
NN	WO200224729-A2.			
PD	28-MAR-2002			
XX				

Query Match	Similarity	72.8%	Score 1006.5	DB 5	Length 273
Best Local	Similarity 71.7%	Pred. No. 1.1e-84			
Matches 198	Conservative 26	Mismatches 49	Indels 3	Gaps 1	
DB	1	MNPFKINGICALASGIALAGCSNOSNEPPAISKTAQTIRYGVWAGPEOAVAEVAGQVAK	60		
QY	1	MMFKINGICALASGIALAGCSNOSNEPPAISKTAQTIRYGVWAGPEOAVAEVAGQVAK	60		
DB	1	MKELQALPAITAIASALVLTGCKEDKKEPEAA--AAPLIKIVGVMSGPEHGVAFETAAKAVAK	57		
QY	61	EKNVLTVELVEFNDYANPNPNSVSGELDANAMQHKPTLEDSQEGKANTVYGVNTVYVP	120		
DB	58	EKVELDQVFEFNDYALPNEVNSGSLDANAMQHKPTLEDSQEGKANTVYGVNTVYVP	117		
QY	121	LASGTSTIKTLELKDQATTAVPNDPSNLARALLILEKQGHILKDKDNTNLFSTLIDVEN	180		
DB	118	LASGTSTIKTLELKDQATTAVPNDPSNLARALLILEKQGHILKDKDNTNLFSTLIDVEN	177		
QY	181	PKKLVIVKVDTSVAAARIDVDLAVVNNNNAAGVGLTASENGVFVEDXDSPPYNIIVARA	240		
DB	178	PKKLVIVKVDTSVAAARIDVDLAVVNNNNAAGVGLTASENGVFVEDXDSPPYNIIVARA	237		
QY	241	DNKDSKAIQDPFVAKYQDEVEAEAKKXQFKQGVIRIGW	276		
DB	238	DNKDSKAIQDPFVAKYQDEVEAEAKKXQFKQGVIRIGW	273		

QY 61 EKKNTLVELVERNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVGTFTYYP 120
 DB 57 EKKGLVQVEVERNDVALPNEAVAKGDLNANQHKPYLEDEKAKNLNLTIVGNTFYYP 116
 QY 121 LAGYSTIKITLNEKDGATIAVPNDPSNARALILEKOGILKLDKNTNLFSTTLDIYEN 180
 DB 117 LAGYSKKIKNVNELDQGAQVVPNDPTNGRALIILEKOGILKLDKNTNLTSTVDIYEN 176
 QY 181 PKKLVYKEVDTSVAPARLDDVLA VNNYNAQVGLTASNGVFEVDKSPYNTIIVARA 240
 DB 177 PKKLTNTVDTISVAPARLDDVLA VNNYNAQVGLTASNGVFEVDKSPYNTIIVART 236
 QY 241 DNKSKAIDPFYKAYOTDEVEAEAKKOPKDGVIKGM 276
 DB 237 DNKSKAVDPFYKSYQTEBEVYQEAQKHFDGVVAKGM 272

RESULT 9
 ABU39463
 ID ABU39463 standard; protein; 276 AA.
 XX
 AC ABU39463;

XX 19-JUN-2003 (first entry)
 XX
 DT Protein encoded by Prokaryotic essential gene #24990.
 XX
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Pasteurella multocida.
 OS

XX WO200277183-A2.
 PN

XX 03-OCT-2002.
 PD

XX 21-MAR-2002; 2002MO-US009107.
 PF

XX 21-MAR-2001; 2001US-00815242.
 PR

XX 06-SEP-2001; 2001US-00948993.
 PR

XX 25-OCT-2001; 2001US-0342923P.
 PR

XX 08-FEB-2002; 2002US-00072851.
 PR

XX 06-MAR-2002; 2002US-0362699P.
 PR

XX (ELIT-) ELITRA PHARM INC.
 PA

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI. 2003-029926/02.
 DR N-PSDB; ACA43333.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PI isolate candidate molecules for rational drug discovery programs.
 XX

XX Claim 25; SEQ ID NO 67387; 1766bp; English.
 PS

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 CC
 CC

SQ Sequence 276 AA;

Query Match 64.2%; Score 888; DB 6; Length 276;
 Best local similarity 60.2%; Pred. No. 1e-73;
 Matches 168; Conservative 55; Mismatches 50; Indels 6; Gaps 3;

QY 1 MNEFKINGICALAGIALAGCSNQSNEPAIS--KTAQTIKVGWAGPECAVAEYAGQ 57
 DB 1 MKLTKLFGLATVLSAVVALAGC--KDKPMAAAPQEPAPARKLTGVWTGAQVTEVAAK 58

QY 58 VAKEXKNTLVELVERNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVGTFTYYP 117
 DB 59 IAKEXKNTLVELVERNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVGTFTYYP 117

QY 118 VYPLAGYSTIKITLNEKDGATIAVPNDPSNARALILEKOGILKLDKNTNLFSTTLDI 177
 DB 118 VFPIAASKKIKNVNELDQGAQVVPNDPTNGRALIILEKOGILKLDKNTNLTSTVDIYEN 177

QY 178 VENPKLVYKEVDTSVAPARLDDVLA VNNYNAQVGLTASNGVFEVDKSPYNTIIVARA 237
 DB 178 IENPKLVYKEVDTSVAPARLDDVLA VNNYNAQVGLTASNGVFEVDKSPYNTIIVARA 237

QY 238 ABADNKSRAIDPFYKAYOTDEVEAEAKKOPKDGVIKGM 276
 DB 238 SREDNKNBAIKDPFYKAYOTDEVEAEAKKOPKDGVIKGM 276

RESULT 10
 ABU48271
 ID ABU48271 standard; protein; 271 AA.
 XX
 AC ABU48271;

XX 19-JUN-2003 (first entry)
 XX
 DT Protein encoded by Prokaryotic essential gene #3798.
 XX
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salomomella typh.
 OS

XX WO200277183-A2.
 PN

XX 03-OCT-2002.
 PD

XX 21-MAR-2002; 2002MO-US009107.
 PF

XX 21-MAR-2001; 2001US-00815242.
 PR

XX 06-SEP-2001; 2001US-00948993.
 PR

XX 25-OCT-2001; 2001US-0342923P.
 PR

XX 08-FEB-2002; 2002US-00072851.
 PR

XX 06-MAR-2002; 2002US-0362699P.
 PR

XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI

PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; AKA52141.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 76195; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 271 AA;
XX
XX Query Match 57.4%; Score 793.5; DB 6; Length 271;
XX Best Local Similarity 56.5%; Pred. No. 5,8e-65;
XX Matches 156; Conservative 44; Mismatches 65; Indels 11; Gaps 3;
QY 3 FGTGNCICATASGIALAGCSNOSNEPRAISKTAOTIKYGVAMGPBOVAEVAEYAGVAKK 62
DB 5 FKTFAVAGALIGSLALAGCGODEKDP-----NHIKGVAVGAEQVAVAEVAKK 56
QY 63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKSOEKGNNLVIYVNTFYVPLA 122
DB 57 YGLDVELVTFNDYVLPREALSKGIDIDANAFQHKPYIDQOKRGY-KLVAVGNTFYVPLA 115
QY 123 GYSTKIKTINELKGATIAVPNDPSNLARALILBKQGLIKLKDNTNLFSTLIDIVENPK 182
DB 116 GYSKRIKSIDELKDSQVAVPNDPTNLGRSLLLQVAGLIKDKDGSLLPTSLDIVENPK 175
QY 183 KLVIRKVDVSVAARAIDD--VDLAAYNNNYAGOVGTASBENFVEDKDSPPVNIIVARA 240
DB 176 NKTIVLELAPQLFRSDDAQIALAVNTTYSQIGLTPAKDGLFVEDKDSPPVNIIVIRE 235
QY 241 DNKDSRAIODFYVAYOTDEVEAEAKKQFQGVYIKGW 276
DB 236 DNKDAENVKKFVAYQSDVEYEAANKVFNGAVKGW 271
RESULT 11
AAU34457
ID AAU34457 standard; protein; 271 AA.

XX AAU34457;
XX AC
XX 14-FEB-2002 (first entry)
XX DT
XX DE E. coli cellular proliferation protein #38.
XX DE
XX Antisense, prokaryotic cellular proliferation protein; antibiotic;
XX antibiotic; drug design.
XX KM
XX Escherichia coli.
XX OS
XX NC0200170955-A2.
XX PN
XX 27-SEP-2001.
XX PD
XX 21-MAR-2001; 2001MO-US009180.
XX PF
XX 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0236252P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX PR
XX (ELIT-) ELITRA PHARM INC.
XX PA
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
XX PI Yamamoto RT, Xu HH;
XX N-PSDB; AAS52316.
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS52316.
XX XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 10050; 51pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes
XX themselves and the encoded proteins. The prokaryotes used are *Escherichia*
XX *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
XX *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence represents an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 271 AA;
XX
XX Query Match 56.6%; Score 782.5; DB 4; Length 271;
XX Best Local Similarity 55.1%; Pred. No. 6.1e-64;
XX Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
QY 3 FGTGNCICATASGIALAGCSNOSNEPRAISKTAOTIKYGVAMGPBOVAEVAEYAGVAKK 62
DB 5 FKTFAVAGALIGSLALAGCGODEKDP-----NHIKGVAVGAEQVAVAEVAKK 56
QY 63 YNLTVLVEFNDYAMPNSAVSKGELDANAMQHPYLEKSOEKGNNLVIYVNTFYVPLA 122
DB 57 YGLDVELVTFNDYVLPREALSKGIDIDANAFQHKPYIDQOKRGY-KLVAVGNTFYVPLA 115
QY 123 GYSTKIKTINELKGATIAVPNDPSNLARALILBKQGLIKLKDNTNLFSTLIDIVENPK 182

PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 55626; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 271 AA;
Query Match 56.4%; Score 780.5; DB 6; Length 271;
Best Local Similarity 56.7%; Pred. No. 9.3e-64;
Matches 152; Conservative 43; Mismatches 62; Indels 11; Gaps 3;
QY 11 ALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVAQVAKKXNLTVEVY 70
DB 13 ALGSLALVCGCGDEKDP-----NHIKVGIVGAEQVAEVAQVAKKXNLTVEVY 64
QY 71 EFNDDYAMNSAVSKGELDANAMQHKPYLEKDSQEKLNLTIVGNTFVYPLAGYSTIKXT 130
DB 65 TFDNDYVLENEALSKGDDANAFQHKPYLDDQIKDRGV-KLVGNGNFFVPIAGYSKIKXS 123
QY 131 LNELEKQATIAVNDPSNLTARALILEKQGLIKLKXNTLFFSTLDIVNPKKLVKXVD 190
DB 124 LDELQPSQVAVNDPPTNLSRSLTLQKQGLIKLKQGVGLPVLVDVTENPKKLKVELE 183
QY 191 TSVAARALID--VDLAVVNNNVAQVGLTPASENGVEVEKDSPEYVNIIVARADNDSKAI 248
DB 184 APQLPRSLDDAQIALAVINTTASQIGTFPAKDGIVFEKDSGYVNLITREDDNDAAENV 243
QY 249 QDFVAKYQDVEFAEAKKQKFGVITKGM 276
DB 244 KKEFQAYQSDVEYQAEANKVENGAVKGM 271
RESULT 14
ABU31676
ID ABU31676 standard; protein; 271 AA.
XX
AC ABU31676;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #17203.

XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Klebsiella pneumoniae.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362639P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Onlsen KL, Zyskind JM,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI; 2003-023926/02.
XX
DR N-PSDB; ACA35546.
XX
PS Claim 25; SEQ ID NO 55600; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 271 AA;
Query Match 56.1%; Score 775.5; DB 6; Length 271;
Best Local Similarity 55.2%; Pred. No. 2.7e-63;
Matches 153; Conservative 45; Mismatches 68; Indels 11; Gaps 3;
QY 2 NFGKINGICALAGSIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAEVAQVAKK 61
DB 4 NFKTFPAVGLIGSLALVCGCGDEKDP-----NHIKVGIVGAEQVAEVAQVAKK 55

CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 271 AA;

Query Match 54.8%; Score 757.5; DB 6; Length 271;
Best Local Similarity 53.2%; Pred. No. 1.3e-61;
Matches 148; Conservative 48; Mismatches 71; Indels 11; Gaps 3;

QY 1 MNEFGKINGICALAGSIALAGCSNQSNEPAISKTAQTIKVGMAGPEQVAEVAQVAK 60
DB 3 LKFKSIALVETALIGTLVCGPTEKAP-----NHIKVGIYVAGQVAVAKVAK 54
QY 61 EKXNLVETLVEFNDYAPNSAVSKGELDANAMOHKRYLEDSOEGKLNVLVGNFTFY 120
DB 55 EKXGLDELVELTFNDYVLPNALSKEGIDILNAFQHKRYLDOQIKDRGY-KLVSVGNSFYVP 113
QY 121 LAGYSTIKITLNEKDGATIAVNDPSNLARALILEKQGLIKLKNNTNLFTTLDIVEN 180
DB 114 IAGSKRKISLDELQPSQVALNDPNTNLSRLISLISQVGLIKKQGVGLPTVLDIVEN 173
QY 181 PKKIVKEVDTSVAARAID--VDLAVNNNVAQVGLTSNGVFEVDKSEFYNIIVA 238
DB 174 PKNLKIVELTAPQLPSRLDQOIALAIINTTVASQIGLFPACKGLFVEDKESFYVNLIVA 233
QY 239 RADKDSKATODPFYKAYQTEVEAEAKKQKDFKGVIGWM 276
DB 234 REDKCDANVKKFVQAYQSDVEYDANKAFNGAVAKGM 271

RESULT 17

ABU40467 standard; protein; 271 AA.

AC ABU40467;
DT 19-JUN-2003 (first entry)
DE Protein encoded by Prokaryotic essential gene #25994.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Proteus sp.
XX
XX WO20027183-A2.
XX

PD 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-03429239.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
DR N-PSDB; ACA44337.
DR WPI; 2003-029926/02.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 68391; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 271 AA;

Query Match 54.8%; Score 757.5; DB 6; Length 271;
Best Local Similarity 52.7%; Pred. No. 1.3e-61;
Matches 147; Conservative 49; Mismatches 70; Indels 13; Gaps 4;

QY 1 MNEFGKINGICALAGSIALAGCSNQSNEPAISKTAQTIKVGMAGPEQVAEVAQVAK 60
DB 3 LKFKSLAVETALVETALVCGPTEKAP-----NHIRVGIVISGEGQVAVAKVAK 54
QY 61 EKXNLVETLVEFNDYAPNSAVSKGELDANAMOHKRYLEDSOEGKLNVLVGNFTFY 119
DB 55 EKXGLDELVELTFNDYVLPNALSKEGIDILNAFQHKRYL--DQIKDRNRYITVAGNFTFY 112
QY 120 PLAGYSTIKITLNEKDGATIAVNDPSNLARALILEKQGLIKLKNNTNLFTTLDIVE 179
DB 113 PLAGYSKRTIDIALDPLDGAQVAIPNDPNTNLSRLISLISQVGLIKKQGVGLPTKDIIE 172

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 11656; 511bp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 240 AA;
Query Match 53.8%; Score 743.5; DB 4; Length 240;
Best Local Similarity 60.0%; Pred. No. 2.1e-60;
Matches 144; Conservative 42; Mismatches 51; Indels 3; Gaps 2;
QY 39 IKVGMAGPEQAVNEVAGOVAKKXNTLVEFENDYAMNSAVSKGELDANMOKHPYL 98
DB 2 IKGVYIGABEQVAEVAQKAKKXGDLVELVTNDVLTNEALSKGIDIVNFKHPYL 61
QY 99 EKDSQEKGLNNLVIVGNTFVYPLAGYSTKTKTNEKDGATVAVNDPSLALILILEK 158
DB 62 DQQIKDNGY-KLVAVGNTFVYPIAGYSKTKSLDELQPSQIIVNDPTMLGRSLILLOQ 120
QY 159 QGLTKLMDNTMLFSTTLDIYENPKKVIKEDTSVAAARA:DD--VDAVANNVYAGQVGL 216
DB 121 VGLTKLMDQVGLFTSLDIYENPKKVIKEDTSVAAARA:DD--VDAVANNVYAGQVGL 180
QY 217 TASNGCVFVEDKDSPPYNIIVARADNKSIAIDPFVAKAYOTDEVEAKKKQFQGVYIKW 276
DB 181 TPADGIFVEGKESPPYNIIVARADNKAENVKFFVQAYQSDVEYBAANKI FNGCAVKGW 240

RESULT 20
ABU19780
ID ABU19780 standard; protein; 295 AA.
AC ABU19780;
XX
XX 19-JUN-2003 (first entry)
DE Protein encoded by Prokaryotic essential gene #5307.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS *Borrelia cepacia*.
XX
XX WO200277183-A2.
PD 03-OCT-2002.
PF 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948593.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX

PA (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA23650.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
PS Claim 25; SEQ ID NO 47704; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid, (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism's activity; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 295 AA;
Query Match 40.6%; Score 562; DB 6; Length 295;
Best Local Similarity 44.5%; Pred. No. 1.9e-43;
Matches 118; Conservative 47; Mismatches 88; Indels 12; Gaps 3;
QY 12 LASGIALAGCSNQSNEPAISKTAOTIKVGMAGPEQAVNEVAGOVAKKXNTLVE 71
DB 37 LATGAALLLSGAHED-----KVIKGTAVAGPDEFWQVQKAKKXGELNVAVIE 88
QY 72 FNDYAMPNSAVSKGELDANAMQHPYLEKDSQEKGLNNLVIVGNTFVYPLAGYSTKTKTL 131
DB 89 FNDYQPPNALDSDGLDANSFOHOPYLDSQVCKRGY-KIVSAGLTYSPIGVSKKPKSL 147
QY 132 NELKDGATVAVNDPSLALILILEKQGLTKLMDNTML--FSTTLDIYENPKKVIK 188
DB 148 KDLPGAGKLAIVNDPSNENRALLLLTQGVTKLAKAGCTGANNATVDDIAPNPKKLISE 207
QY 189 VDTSVAAARAIIDVDLAVANNVYAGQVGLTASNGCVFVEDKDSPPYNIIVARADNKSIAI 248
DB 208 LDAQGLPRLVLDVDAVINTVYALANLQPTKDAIALESLSLTPYANLIIAVAKKDKQFPWV 267
QY 249 QDFVAKAYOTDEVEAKKKQFQGVYI 273
DB 268 KDLVAYQSPVKERIKKQFGSMV 292

RESULT 21

ABU23447
ID ABU23447 standard; protein; 262 AA.
XX
AC ABU23447;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #8974.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bordetella pertussis.
XX
PN W0200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA27317.
PT
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX
XX Claim 25; SEQ ID NO 51371; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 262 AA;

Query Match 39.8%; Score 551; DB 6; Length 262;

Best Local Similarity 43.1%; Pred. No. 1.6e-42; Matches 118; Conservative 52; Mismatches 86; Indels 18; Gaps 5;

QY 1 MNFGKINGICALASGIALAGCSNQSNBPAAISKTAAGTIKVGWAGPQWAEVA-GQVA 59
DB 3 MNFVR---SALLASFLLAGGAQ-----AEKLVGATQVPAHEILEVKKPALA 48
QY 60 KEKNLITVELVFPNDYAMNSAVSGEDANAMQHKPFLKDSQSKGLNNLYIGENTPY 119
DB 49 KGVVELDIKVV-FTTYVQPNQLADKQDANFQHPILDTFNKR-KTNLVSVELVAHE 105
QY 120 PLAGYSTKIKTINEDGATIAVPNDPSNLAAALILEKQGLIKLKDWNLFSTLLDIVE 179
DB 106 PFGYISKTKISLAELKDGATIAIPNDPNSGALLLLKQGLKXKDPNSIVATIDIAE 165
QY 180 NPKLVKEVDTSVAPAIIDVDIAVNNNAGVQGLTASENGVEDEKDSFYNIIVAR 239
DB 166 NPKLKEFLELAAMPDRSFDDLDLINTVYALAEAGLVPTRDALFIEGADSEYANLVAA 225
QY 240 ADNKDSKAIQDFPKAYQGTDEVEAEAKQKQGI 273
DB 226 PNKDPAVAKLVNLAHSEAVKFTIEKIGAVV 259

RESULT 22

ABU22451
ID ABU22451 standard; protein; 272 AA.
XX
AC ABU22451;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #7978.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Burkholderia mallei.
XX
PN W0200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA26321.
PT
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 50375; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 272 AA;

Query Match 39.7%; Score 549.5; DB 6; Length 272;
Best Local Similarity 42.5%; Pred. No. 2.4e-42;
Matches 114; Conservative 52; Mismatches 91; Indels 11; Gaps 3;

QY 9 ICALASGATAGCSQNSNEPAIKTAQITKVGWMAPEQVAEVAQVAKKXNTVE 68
DB 10 VSAVAAGAAASVS-----VGAQADQKVIKGTAVAGPDQVQVQKAKKQGLDVK 62
QY 69 LVEFNDVYQVPAALDSDGLDANSFQHQPYLDSQVQKQGY-KIVASGLTISPIGVSKKF 121
DB 63 VLEFNDVYQVPAALDSDGLDANSFQHQPYLDSQVQKQGY-KIVASGLTISPIGVSKKF 121
QY 129 KTLNELKQGAITAVPNPDSNLRALILILEKQGLIKLKDNTML---FSTTLDIVENPKLV 155
DB 122 KSLVELPAGAKVVALPNPDSNLRALILILEKQGLIKLKDNTML---FSTTLDIVENPKLV 155
QY 186 IKEYDTSVAARADVDVLAANNVYAGVGLTASENGVEFEDKDSFYVNIIVARADNKDS 245
DB 182 LTELDAAGLQPRVLSVDAAVINTVYLAANLQPKDAIALESLTSPYANLIAVARAKDQ 241
QY 246 KAIQDFVAKYQTDVEVEAKKQFQDGYI 273
DB 242 FVWKGLVAKYQSPVKERITQFQGSWV 269

RESULT 23
ABU33203
ID ABU33203 standard; protein: 256 AA.
XX
XX ABU33203;
AC 19-JUN-2003 (first entry)
DT 19-JUN-2003 (first entry)
DE Protein encoded by prokaryotic essential gene #18730.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Legionella pneumophila.
XX
XX WO200277183-A2.
PN 03-OCT-2002.
PD 21-MAR-2002; 2002WO-US009107.
PF 21-MAR-2001; 2001US-00815242.
XX
XX

PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0352699P.
XX (EIT-) ELITRA PHARM INC.
PA Wang L, Zmudis C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall L, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA37073.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 61127; 1766bp; English.
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 621 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 256 AA;

Query Match 39.4%; Score 545; DB 6; Length 256;
Best Local Similarity 45.7%; Pred. No. 5.7e-42;
Matches 118; Conservative 37; Mismatches 91; Indels 12; Gaps 3;

QY 16 IALAGCSQNSNEPAIKTAQITKVGWMAPEQVAEVAQVAKKXNTVELEFNDY 75
DB 11 ISLVACS-----SKPAPVTLVIGTLAGETELTETAKOVAKEKGLIKIVLEFNDY 61
QY 76 AMFNSAVSKGELDANAMQKRYLEKDSQKGLNIVYGNFVYVPLAGYKRTIKTINELK 135
DB 62 NLPNEALDQGSIDAVYQHLPLKKAIIISHG-DIQALIGRTFVYVPMGYSKRTILSLP 120
QY 136 DGATIAVNDPNSNLARALILILEKQGLIKLKDNTMLFSTTLDIVENPKLVKEYDTSVA 195
DB 121 ENGIIAIVNDPENEMRAFLLEKKAHLITLKNTTN--SGIQIENSPPKQFKREIDAQLP 178
QY 196 RAIDVDVLAANNVYAGVGLTASENGVEFEDKDSFYVNIIVARADNKDSKAIQDFVAKY 255
DB 179 RVLPPVDAVINTVYLAANLQPKDAIALESLTSPYANLIAVARAKDQ 238

QY 256 QTDVEAEAKKQFEDGVI 273
DB 239 NSXEVKEKXNLFGEPAI 256

RESULT 24

AB021332
ID AS021332 standard; protein; 268 AA.
XX
AC AB021332;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #6859.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Burkholderia fungorum.
XX
PN W0200277183-A2.
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
DR N-PSDB; ACA25202.
XX
PF 2003-029926/02.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 49256; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the fastest prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 268 AA;

Query Match 39.0%; Score 540; DB 6; Length 268;
Best Local Similarity 44.6%; Pred. No. 1,8e-41;
Matches 108; Conservative 49; Mismatches 83; Indels 2; Gaps 2;

QY 35 AAQTIKGVMAQPEQAVAEVAGVAKKYNLTVELVFNDFYAMNSAIVSKGLDANAMCH 94
DB 29 AEDTIKVGVTGSPHQAQMEVVKTYA-AKNGLNIVKVESDIYQVPAALAGSDLDNSYQH 87
QY 95 KPYLEKDSQEKGLNNLVVGNTPFVYPLAGYSTIKITINELKDGATTIAPNDPSMLARALI 154
DB 88 DPLYLAQVYKDRGY-KLIRVADTVTYPMGIYSKVKYSLAEILOGAKIAPNDPTNGRALL 146
QY 155 LLEKQGLIKLNDNTNLFSTTIDIVENPKKVIKVEYDVSVAARAIDDVDLAVNNNYAGQV 214
DB 147 LLKQGLIKLRADAGLAKTPDIDVNPCKLIVELDAAQIPRSINDVDAAININFAEBA 206
QY 215 GLTASENGVPVEDKDSFYVNIIVARADKDSKAIODFVKAQTDVEAEAKKQFEDGVIK 274
DB 207 GLKPKQDAIAIEDPKGPIYVNIARADRNKFWAKLVAAHSPFVKQFVSKFGGSVIT 266
QY 275 GW 276
DB 267 AW 268

RESULT 25

AB021415
ID AB021415 standard; protein; 270 AA.
XX
AC AB021415;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #6942.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Burkholderia fungorum.
XX
PN W0200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
DR N-PSDB; ACA25285.
XX
PF 2003-029926/02.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 49339; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression


```
QY 189 VDTSVARAIDVDVLA VVNNNAGVGLTPASENGVEVEDKSPYNNIIVARADNRKSKAI 248
DB 178 LDAAGQPRSLDDTDASAVNNNFALFAGLDPSKALVRESAESPYANVLVVRBQDKDRADL 237
QY 249 QDFVYKAYQTDVEVEAEAKKQFKDGVIKGW 276
DB 238 RKLVSIVQSAFVREFLIGKTKYGA VAAW 265

RESULT 27
ABB53625
ID ABB53625 standard; protein; 286 AA.
AC ABB53625;
XX
XX 29-AUG-2003 (revised)
DT 16-MAY-2002 (first entry)
XX
DE Lactococcus lactis protein plpB.
XX
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS Lactococcus lactis; IL1403.
XX
XX FR2807446-A1.
XX
XX 12-OCT-2001.
XX
XX 11-APR-2000; 2000PR-00004630.
XX
XX 11-APR-2000; 2000PR-00004630.
XX
XX 11-APR-2000; 2000PR-00004630.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
XX Bolocline A, Sorokine A, Renault P, Ehrlich SD;
XX
XX WPI; 2002-043418/06.
XX
XX New nucleotide sequence useful in the identification or Lactococcus
XX
XX lactis and related species.
XX
XX Claim 6; SEQ ID NO 327; 2504pp; French.
XX
XX The present invention is related to a Lactococcus lactis nucleotide
XX
XX sequence (AA9521) and related proteins (AB53300-AB55621). The nucleic
XX
XX acid sequence is useful in the detection and/or amplification of nucleic
XX
XX acid sequence, particularly to identify Lactococcus lactis or related
XX
XX species. The proteins of the invention are useful for the biosynthesis or
XX
XX biodegradation of a composition of interest. The invention helps research
XX
XX in lactic bacteria, particularly useful in the production of yogurt and
XX
XX cheese. Note: The sequence data for this patent is based on equivalent
XX
XX patent WO20017734 (published 16-OCT-2001) which is available in
XX
XX electronic format directly from WIPO at
XX
XX ftp.wipo.int/pub/published_pot_sequences. (Updated on 29-AUG-2003 to
XX
XX standardise OS field)
XX
XX Sequence 286 AA;
XX
XX Query Match 36.8%; Score 509.5; DB 5; Length 286;
XX
XX Best Local Similarity 40.5%; Pred. No. 1.3e-38;
XX
XX Matches 111; Conservative 58; Mismatches 90; Indels 15; Gaps 8;
```

```
DB 129 QLNKISDIKEGTTIAVPNDASNESRALVYLKAGLIKLDVSGOTLATVKDITSPKXLI 188
QY 187 KEVDTSVARAIDVDVLA VVNNNAGVGLTPASENGVEVE--DDDS--PYNNIIVARADNR 243
DB 189 KELDASQTAARALDSVDAAVNNNVAVTGRLKSD-AIFTEPVNRDSCQMIIIVANKDE 247
QY 244 DSKAIQDFVYKAYQTDVEVEAEAKKQFKD-GVIKGW 276
DB 248 NNTVYKDVYKAYETEATKTKIAKAVPDKSTIPAW 281

RESULT 28
ABU39768
ID ABU39768 standard; protein; 261 AA.
AC ABU39768;
XX
XX 19-JUN-2003 (first entry)
DT
XX
XX Protein encoded by Prokaryotic essential gene #25295.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Pseudomonas putida.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW,
XX
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX
XX N-PGDB; ACA43638.
XX
XX Claim 25; SEQ ID NO 67692; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX
XX the 623 antisense sequences given in the specification where expression
XX
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX
XX encoding a polypeptide whose expression is inhibited by the antisense
XX
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX
XX polypeptide or its fragment whose expression is inhibited by the
XX
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX
XX proliferation or the activity of a gene in an operon required for
XX
XX proliferation; (7) identifying a compound that influences the activity of
XX
XX the gene product or that has an activity against a biological pathway
XX
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX
XX identifying a gene required for cellular proliferation or the biological
XX
XX pathway in which a proliferation-required gene or its gene product lies
XX
XX or a gene on which the test compound that inhibits proliferation of an
XX
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX
XX compound's activity; (11) a culture comprising strains in which the gene
XX
XX product is overexpressed or underexpressed; (12) determining the extent
XX
XX to which each of the strains is present in a culture or collection of
XX
XX strains; or (13) identifying the target of a compound that inhibits the
XX
XX proliferation of an organism. The antisense nucleic acids are useful for
```

identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 261 AA:

Query Match 36.6%; Score 506.5; DB 6; Length 261;
Best Local Similarity 41.3%; Pred. No. 2.2e-38;
Matches 109; Conservative 52; Mismatches 88; Indels 15; Gaps 4;

11 ALASGIALAGCNSQNSNPAAISKTAQTIKVGWAGBEQAVAR-VAGQVAKKYNLTVEL 69
9 ALAAASFSGLA-----AAAEKLVVATPPVPHAEILBLIFPTLAKGVNDLQIV 57
QY 70 VEFNDVAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNNTLVGNTFYVPLAGYSTKIK 129
DB 58 --FDYVQPVVQVQDKRLDANVFQTLPLYQNFNEGKTHLETYVG--VHVEPFGYSEKVK 114
QY 130 TINEIKGATITVNDNSNLARALILEKGLIKLKDNTNLFSTLTDIYENPKKIVIKEV 189
DB 115 ALSELKGGATVVALPNEGSSNGRALLLQKAGLITLKDPPKVALTPKDIENPKLKFREL 174
QY 190 DTSAARAIDVDVLAIVNNVYAGVGLTASENGVEVEDKSPYVNIIVARADNKSKAIO 249
DB 175 ESAMLPVPLDQVLDIMNTNYALEAGLNPAAKDALVIEGADSPVNPFLVARPDKDSEALQ 234
QY 250 DFVKAQYTTDEVKAEAKQFQDGYI 273
DB 235 KLAQALTSPEVKAFIAKKYQAVL 258

RESULT 29

AAU36517 standard; protein; 260 AA.

AAU36517;

14-FEB-2002 (first entry)

Pseudomonas aeruginosa cellular proliferation protein #507.

Antisense; prokaryotic cellular proliferation protein; antibiotic;
antibacterial; drug design.

Pseudomonas aeruginosa.

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US009180.

21-MAR-2000; 2000US-0191078P.

23-MAY-2000; 2000US-0206848P.

26-MAY-2000; 2000US-0207727P.

23-OCT-2000; 2000US-0242578P.

27-NOV-2000; 2000US-0253625P.

22-DEC-2000; 2000US-0257931P.

16-FEB-2001; 2001US-0269308P.

(ELIT-) ELITRA PHARM INC.

Haseelbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
Yamamoto RT, Xu HH,
WPI; 2001-611495/70.
DR N-PSDB; AAS54376.

New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.
Example 3; SEQ ID NO 12110; 511p; English.

The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are *Escherichia*
coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
Pseudomonas aeruginosa and *Enterococcus faecalis*. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence represents an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
[ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 260 AA:

Query Match 36.3%; Score 502.5; DB 4; Length 260;
Best Local Similarity 42.6%; Pred. No. 5e-38;
Matches 106; Conservative 52; Mismatches 84; Indels 7; Gaps 3;

29 AAIKTAQTIKVGWAGBEQAVAEVAGV-----AKEKNVLTVELVEFNDVAMPNSAVSK 84
12 AALGITAQAASLTVATPPVPHAEILNIVKYKPLAKGVLDLKK--EFTYVQPVVQVSE 69
QY 85 GELDANAMQHKPYLEKDSQEKLNNTLVGNTFYVPLAGYSTKIKTINELKDGATIVEN 144
DB 70 KRLDANFQHPQYIDENKAKG--TDLVAVTGVIHEPLGAVSKKKKDELPSGATVIVP 128
QY 145 DPSNLARALILEKGLIKLKDNTNLFSTLTDIYENPKKIVIKEVNDVSAARAIDVDLA 204
DB 129 DATNGRALLLDDRAGVAKDKNSITATPKDIDVNDKNIKIRLEBAATTPTVITQVDNA 188
QY 205 VVNNVYAGVGLTASENGVEVEDKSPYVNIIVARADNKSKALQDFVKAQYQDEVKAE 264
DB 189 LINTNYALEAKLNTKDALIEGSDSPYVNIIVARPDNKSDAMQKLAKLHSAEIKOFI 248
QY 265 KKQFQDGYI 273
DB 249 QEKYKGAIV 257

RESULT 30

ABU38920 standard; protein; 260 AA.

ABU38920;

19-JUN-2003 (first entry)

Protein encoded by prokaryotic essential gene #24447.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Pseudomonas aeruginosa.

WO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.


```
Db      7 LTLASVY-----SVSASAEXLIVAATPIPAHELLEIKPTLAEGVDLQI 53
Qy      68 ELVEFNDYAMPNSAVSGEIDANAMQHPYLEKDSQEKGLNNLVIGNTFVYPLAGYSTK 127
Db      54 KV--FTDYQPNVAVAEKRLDANVFQTLPLENPNKSGKGNLTAVG-VHVEPFGYSRK 110
Qy      128 IKTINELKDGATTAVNDPSNLARALILEKQGLIKDKNTNLFSTLLDIVENPKLVIR 187
Db      111 YKSLAELPDGATVAIPNEGNSGRALLILOKAGILTKPKNNMLATPKDIAENPKNLKPK 170
Qy      188 EVDTSVAAARAIDVDVLAANNVAGVGLTASENGVEDEKDSFYNIIVARADNKDSDA 247
Db      171 ELBSALLPRVLDQVDLINTNTALAEKLNPAKDALVLEBRDSPYNYVVARPDNKDSDA 230
Qy      248 IQDFVKAAYQTFDEVEAEAKKQPKDQVI 273
Db      231 LKKLSALTSPVAVAFIEKKYAGAVV 256

RESULT 32
ABU15545
ID      ABU15545 standard; protein; 259 AA.
XX
AC      ABU15545;
XX
DT      19-JUN-2003 (first entry)
XX
DE      Protein encoded by Prokaryotic essential gene #1072.
XX
KW      Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS      Pseudomonas aeruginosa.
XX
PN      MO200277183-A2.
XX
PD      03-OCT-2002.
XX
PF      21-MAR-2002; 2002MO-US009107.
XX
PR      21-MAR-2001; 2001US-00815242.
PR      06-SEP-2001; 2001US-00948893.
PR      25-OCT-2001; 2001US-0342923P.
PR      08-FEB-2002; 2002US-00072851.
PR      06-MAR-2002; 2002US-0362699P.
XX
XX
PA      (ELIT-) ELITRA PHARM INC.
PI      Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind UW,
PI      Wall D, Trawick UD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX      WPI; 2003-029926/02.
DR      N-PSDB; ACA19415.
XX
XX      Claim 25; SEQ ID NO 43469; 176bp; English.
XX
XX      The invention relates to an isolated nucleic acid comprising any one of
XX      the 6213 antisense sequences given in the specification where expression
XX      of the nucleic acid inhibits proliferation of a cell. Also included are:
XX      (1) a vector comprising a promoter operably linked to the nucleic acid
XX      encoding a polypeptide whose expression is inhibited by the antisense
XX      nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX      polypeptide or its fragment whose expression is inhibited by the
XX      antisense nucleic acid; (4) an antibody capable of specifically binding
XX      the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX      proliferation or the activity of a gene in an operon required for
XX      proliferation; (7) identifying a compound that influences the activity of
XX      the gene product or that has an activity against a biological pathway
XX      required for proliferation, or that inhibits cellular proliferation; (8)
```

```
CC      identifying a gene required for cellular proliferation or the biological
CC      pathway in which a proliferation-required gene or its gene product lies
CC      or a gene on which the test compound that inhibits proliferation of an
CC      organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC      compound's activity; (11) a culture comprising strains in which the gene
CC      product is overexpressed or underexpressed; (12) determining the extent
CC      to which each of the strains is present in a culture or collection of
CC      strains; or (13) identifying the target of a compound that inhibits the
CC      proliferation of an organism. The antisense nucleic acids are useful for
CC      identifying proteins or screening for homologous nucleic acids required
CC      for cellular proliferation to isolate candidate molecules for rational
CC      drug discovery programs, or for screening homologous nucleic acids
CC      required for proliferation in cells other than S. aureus, S. typhimurium,
CC      K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC      the target prokaryotic essential genes. Note: The sequence data for this
CC      patent did not form part of the printed specification, but was obtained
CC      in electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 259 AA;
XX
Query Match      36.0%; Score 497.5; DB 6; Length 259;
Best Local Similarity 41.4%; Pred. No. 1.5e-37;
Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;

Qy      9 ICALAGGIALAGCGNNEPAAISKTAQTIKYGVNAGPEQVAE-VAGQVAKKYNLTV 67
Db      7 LTLASVY-----SVSASAEXLIVAATPIPAHELLEIKPTLAEGVDLQI 53
Qy      68 ELVEFNDYAMPNSAVSGEIDANAMQHPYLEKDSQEKGLNNLVIGNTFVYPLAGYSTK 127
Db      54 KV--FTDYQPNVAVAEKRLDANVFQTLPLENPNKSGKGNLTAVG-VHVEPFGYSRK 110
Qy      128 IKTINELKDGATTAVNDPSNLARALILEKQGLIKDKNTNLFSTLLDIVENPKLVIR 187
Db      111 YKSLAELPDGATVAIPNEGNSGRALLILOKAGILTKPKNNMLATPKDIAENPKNLKPK 170
Qy      188 EVDTSVAAARAIDVDVLAANNVAGVGLTASENGVEDEKDSFYNIIVARADNKDSDA 247
Db      171 ELBSALLPRVLDQVDLINTNTALAEKLNPAKDALVLEBRDSPYNYVVARPDNKDSDA 230
Qy      248 IQDFVKAAYQTFDEVEAEAKKQPKDQVI 273
Db      231 LKKLSALTSPVAVAFIEKKYAGAVV 256

RESULT 33
ABU25052
ID      ABU25052 standard; protein; 263 AA.
XX
AC      ABU25052;
XX
DT      19-JUN-2003 (first entry)
XX
DE      Protein encoded by Prokaryotic essential gene #10579.
XX
KW      Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS      Clostridium difficile.
XX
PN      MO200277183-A2.
XX
PD      03-OCT-2002.
XX
PF      21-MAR-2002; 2002MO-US009107.
XX
PR      21-MAR-2001; 2001US-00815242.
PR      06-SEP-2001; 2001US-00948893.
PR      25-OCT-2001; 2001US-0342923P.
PR      08-FEB-2002; 2002US-00072851.
PR      06-MAR-2002; 2002US-0362699P.
XX
XX
PA      (ELIT-) ELITRA PHARM INC.
```


XX	ABU41435	standard; protein; 257 AA.
XX	ABU41435,	
XX	19-JUN-2003	(first entry)
XX	Protein encoded by Prokaryotic essential gene #26962.	
XX	Antisense; prokaryotic essential gene; cell proliferation; drug design.	
XX	Pseudomonas syringae.	
XX	WO200277183-A2.	
XX	03-OCT-2002.	
XX	21-MAR-2002; 2002WO-US009107.	
XX	21-MAR-2001; 2001US-00815242.	
XX	06-SEP-2001; 2001US-00948993.	
XX	25-OCT-2001; 2001US-0342923P.	
XX	08-FEB-2002; 2002US-00072851.	
XX	06-MAR-2002; 2002US-0362699P.	
XX	(ELIT-) ELITRA PHARM INC.	
XX	Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zyskind JW;	
XX	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
XX	WPI; 2003-029926/02.	
XX	N-PSDB; ACA45305.	
XX	New antisense nucleic acids, useful for identifying proteins or screening	
XX	for homologous nucleic acids required for cellular proliferation to	
XX	isolate candidate molecules for rational drug discovery programs.	
XX	Claim 25; SEQ ID NO 69359; 1766bp; English.	
XX	The invention relates to an isolated nucleic acid comprising any one of	
XX	the 613 antisense sequences given in the specification where expression	
XX	of the nucleic acid inhibits proliferation of a cell. Also included are:	
XX	(1) a vector comprising a promoter operably linked to the nucleic acid	
XX	encoding a polypeptide whose expression is inhibited by the antisense	
XX	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
XX	polypeptide or its fragment whose expression is inhibited by the	
XX	antisense nucleic acid; (4) an antibody capable of specifically binding	

	Query Match	Similarity	Best Local	Matches	Conservative	Mismatches	Indels	Gaps
	35.2%	Score 487;	DB 6;	Length 257;				
	42.9%	Pred. No. 1,4e-36;						
	105;	45;	91;	4;	3			
Qy	30	AISKTAQTIKVGVAAGEQAVAR-VAGQVAKKEKINLTVELVFNDYAMPNSAVSGEID	88					
Db	13	ALSAQNETLTLVAAGVAPHAILEFVFPFLAKEGVDLSEICV--FNDYIOPNNOVSGSRID	70					
Qy	89	ANAMCHKEPLEKDSQEKANLNLVIGMTFVYPLAGVSTKIKITNELKDGTIAPVDPEN	148					
Db	71	ANFPQAPLDEIRYNGKG-TDLVAAYAKVHEPFGAYSDRKLDLPLAGANVALPRDATT	129					
Qy	149	LARALLLEKQGLIKLKDNTNLFSTLIDVENPKKIVIKEVDTSVARAIDVYDLAVNN	208					
Db	130	BEGALLILAKAGLITLKDPNLTSEKSPDVANNPKDLKFELEAATLPVLTQVLDALINT	189					
Qy	209	NVAQGVGLTASNGVFDKDSFPYNTITVARADKDSKALQDPVKAYQDVEAEAKKOF	268					
Db	190	NVALSKAPDPTKDALIEGADSPYANILVTRPNKDSDAIKLVALAQSPVKTPLAEKY	249					
Qy	269	KDGVI 273						
Db	250	KGAIV 254						
RESULT 36								
ID	AA00024							
ID	AA00024	standard; protein; 277 AA.						
XX	AA00024;							
XX	20-APR-1999	(first entry)						
XX	DT							
XX	Enterococcus faecalis	protein EF009.						
DE	Enterococcus faecalis	infection; vaccine; immune response; diagnosis;						
KM	detection; attenuation;	antigenic.						
XX	Enterococcus faecalis.							
OS	WC0850554-A2.							
XX	12-NOV-1998.							
XX	04-MAY-1998;	98WC-US008959.						
XX	06-MAY-1997;	97US-0044031P.						

PR 04-MAY-1998; 98US-00071035.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX
DR WPJ: 2002-425450/45.
DR N-PDSB; ABNS97999.
XX
PT New genes and polypeptides from *Enterococcus faecalis*, useful as vaccines
PT for preventing, treating or attenuating an infection caused by a member
FI of the *Enterococcus* genus in an animal, particularly *E. faecalis*.
PS
XX
Claim 9; Page 35; 255pp; English.

CC The present invention provides the protein and coding sequences of a
CC number of polypeptides from *Enterococcus faecalis*. The proteins can be
CC used as vaccines for preventing or attenuating an infection caused by a
CC member of the *Enterococcus* genus in an animal, particularly *E. faecalis*.
CC The polynucleotide is also useful for preventing or treating *E. faecalis*
CC infection. The present sequence is a protein of the invention
XX
XX
SQ Sequence 277 AA;

Query Match 34.8%; Score 481; DB 5; Length 277;
Best Local Similarity 36.8%; Pred. No. 5,4e-36;
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5

```

QY      3 FGKINQICLACGIALACSCNQSNEPALISXTAQQITKYVMAGPPOQVAEYGVAXEK 62
Db      4 FSLTGLIGIYLA-FLIACGASGS-----VADTKETIYKLGAVGTGKNDWESVDRKKK- 56
QY      63 YNLTVELVENDYAMPNSAVSGEILDANMOHKPYLEKDSQOEKGLNNLYVGNTEVPYLA 122
Db      57 -IDJQVLEFDTYTOPNALAEKEIDINAFQHOIFLNNYKREHG-TELYSIGNTVAAPLG 114
QY      123 GYSTKTKTNELKDGATAYVNDPSNLAPALITLEKQGLKLDNNNLSTLIDIVENPK 182
Db      115 IYANKIKDITIKDQSGEIALIPNDPTNGRALILLOTGLIKVDPACQCLPTYSDITENK 174
QY      183 KLVIKEVDTSVARAIDVDLAVVNNNYAQOVGLTASENGVFE---DKDSPYVNIYAR 239
Db      175 QKITELELATQARALQVDDASVINSQAVDAGYTPBKDAIFLEPNEKAKPYNIVAR 234
QY      240 ADMNKSKALQDFVYKAYOTDEVEAKKQFQDGYIKGH 276
Db      235 EEDQENKLYQKVEBYQOEETTKVIATLSTSGANVPAM 271

```

XX	AC	ABU88271	standard; protein, 277 AA.
XX	AC	ABU88271;	
XX	DT	07-JUL-2003	(first entry)
XX	DE	E. faecalis novel protein #15.	
XX	KM	Vaccine; endocarditis; bacteraemia; urinary tract infection; UTI;	
XX	KW	intraabdominal infection; soft tissue infection; neonatal sepsis.	
XX	OS	Enterococcus faecalis.	
XX	PN	US2003017495-A1.	
XX	PD	23-JAN-2003.	
XX	PF	29-JUL-2002; 2002US-00206576.	
XX	PR	06-MAY-1997; 97US-0044031P.	
XX	PR	16-MAY-1997; 97US-0046655P.	
XX	PR	14-NOV-1997; 97US-0066008P.	

RESULT 39
 ABU29285
 ID ABU29285 standard; protein; 277 AA.
 XX
 XX ABU29285;
 XX
 XX 19-JUN-2003 (first entry)
 XX
 XX Protein encoded by Prokaryotic essential gene #14812.
 DE
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 XX Enterococcus faecalis.
 OS
 XX
 XX MO200277183-A2.
 XX
 XX
 XX 03-OCT-2002.
 XX
 XX
 XX 21-MAR-2002; 2002MO-US009107.
 PF
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 23-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wai D, Treweek JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 XX WPI; 2003-029926/02.
 DR
 XX N-PSDB; ACA31155.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX
 XX Claim 25; SEQ ID NO 57209; 1766bp; English.
 PS
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 XX the 613 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 277 AA;
SQ

Query Match 34.8%; Score 481; DB 6; Length 277;
Best Local Similarity 36.8%; Pred. No. 5.4e-36;
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAAQTIVGVMAGPEQAVAEVAGVAKK 62
DB 4 FSKLIGLIGVLA-FTIAGCAGS-----VKDTKETVTLGVGTGKNDWESVDRKKK- 56
QY 63 YNLTVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNNTVIGNTFVYPLA 122
DB 57 -NIDQLVEFTDYTPNPAALAEKEIDLNAFOHQIFLDVYNKEHG-TKLVSIGNTVNAAPLG 114
QY 123 GYSTRIKTLNELKDGTIAVNDPSNLARALILEKQGLIKLKNNTLFTLLDIENPK 182
DB 115 IYANKLKDITIKIDGGEIAlPNDPTNGRALIILQTAGLIKVDPAKQQLPTVSDITENKR 174
QY 183 KLVKEVDTVAARAIDVDLAVVNNNYAGQVGLTASENGVEFVE---DKDSPYNIIVAR 239
DB 175 QLKITELDATQTAARALQDVDAVINSQAVDAGYTPDKDAIFLEPVNEKAKPYNIIVAR 234
QY 240 ADNKSRAIDPFKAYOTDEVEAEAKQKQFQGVIKGW 276
DB 235 EEDQENKLYQKVEEYQOEETKVIATETSKGANNVPAW 271

RESULT 40
ABU13522

ID ABU13522 standard; protein; 277 AA.

AC ABU13522;

XX 26-FEB-2003 (first entry)

XX Enterococcus faecalis EF040 polypeptide #15.

XX EF040; immunostimulant; antibacterial; gene mapping.

XX Enterococcus faecalis.

XX US6448043-B1.

XX 10-SEP-2002.

XX 04-MAY-1998; 98US-00071035.

XX 06-MAY-1997; 97US-0044031P.

XX 16-MAY-1997; 97US-004655P.

XX 14-NOV-1997; 97US-0066009P.

XX 14-NOV-1997; 97US-0066009P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Choi GH, Bailey C, Hromockyj A, Kunach CA;

XX WPI; 2003-089120/08.

XX N-PSDB; ABX61569.

XX Example 1; Col 55-56; 146pp; English.

CC The invention relates to polynucleotide fragments of a gene from
CC Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
CC polypeptides are useful in detecting E. faecalis, as epitope tags, as
CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel
CC filtration columns, in generating antibodies that specifically bind to
CC the E. faecalis polypeptides, in generating an immune response against E.
CC faecalis and other Enterococcus species and as vaccines against other

CC bacterial genera. The polynucleotides are useful as probes for gene
CC mapping and for identifying E. faecalis in biological samples. Sequences
CC ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:
CC The sequence data for this patent can also be obtained from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 277 AA;
SQ

Query Match 34.8%; Score 481; DB 6; Length 277;
Best Local Similarity 36.8%; Pred. No. 5.4e-36;
Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAAQTIVGVMAGPEQAVAEVAGVAKK 62
DB 4 FSKLIGLIGVLA-FTIAGCAGS-----VKDTKETVTLGVGTGKNDWESVDRKKK- 56
QY 63 YNLTVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNNTVIGNTFVYPLA 122
DB 57 -NIDQLVEFTDYTPNPAALAEKEIDLNAFOHQIFLDVYNKEHG-TKLVSIGNTVNAAPLG 114
QY 123 GYSTRIKTLNELKDGTIAVNDPSNLARALILEKQGLIKLKNNTLFTLLDIENPK 182
DB 115 IYANKLKDITIKIDGGEIAlPNDPTNGRALIILQTAGLIKVDPAKQQLPTVSDITENKR 174
QY 183 KLVKEVDTVAARAIDVDLAVVNNNYAGQVGLTASENGVEFVE---DKDSPYNIIVAR 239
DB 175 QLKITELDATQTAARALQDVDAVINSQAVDAGYTPDKDAIFLEPVNEKAKPYNIIVAR 234
QY 240 ADNKSRAIDPFKAYOTDEVEAEAKQKQFQGVIKGW 276
DB 235 EEDQENKLYQKVEEYQOEETKVIATETSKGANNVPAW 271

Search completed: June 16, 2004, 11:06:40
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:00:51 ; Search time 18 Seconds
(without alignments)
798.409 Million cell updates/sec

Title: US-10-018-672-2
Sequence: 1 MNFKINGICALLAGSLALAG.....TDEVEBAKKQFKGVINGW 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1012.5	73.2	277	1	PLPA_PASHA
2	1006.5	72.8	273	1	METQ_PASHA
3	888	64.2	276	1	METQ_PASHA
4	793.5	57.4	271	1	METQ_PASHA
5	792.5	57.3	271	1	METQ_PASHA
6	791.5	57.2	263	1	METQ_PASHA
7	791.5	57.2	271	1	METQ_PASHA
8	782.5	56.6	271	1	METQ_PASHA
9	773	55.9	276	1	PLPB_PASHA
10	757.5	54.8	271	1	METQ_PASHA
11	751.5	54.3	269	1	METQ_PASHA
12	691.5	50.0	272	1	MLPA_PASHA
13	686.5	49.6	272	1	MLPA_PASHA
14	416.5	30.1	268	1	TPP2_PASHA
15	286	21.4	263	1	YHCO_PASHA
16	112.5	8.1	267	1	GRF2_PASHA
17	106	7.7	1592	1	GRF2_PASHA
18	104.5	7.6	991	1	SCA4_PASHA
19	103.5	7.5	543	1	CH60_PASHA
20	103	7.4	1597	1	GRF1_PASHA
21	99.5	7.2	465	1	APPA_PASHA
22	99	7.2	814	1	SLA1_PASHA
23	99	7.1	928	1	HXA2_PASHA
24	98.5	7.1	459	1	TRWE_PASHA
25	98	7.1	469	1	YH89_PASHA
26	97.5	7.0	306	1	OPBC_PASHA
27	97.5	7.0	1164	1	YH89_PASHA
28	97	7.0	1396	1	YH89_PASHA
29	96.5	7.0	339	1	PUR3_PASHA
30	96.5	7.0	2144	1	GLT1_PASHA
31	96	6.9	382	1	VATC_PASHA
32	96	6.9	428	1	ARO4_PASHA
33	96	6.9	525	1	SPI_PASHA

ALIGNMENTS

34	96	6.9	546	1	CH60_AZOVI	08gb4 azobacter
35	96	6.9	1729	1	RBP5_YEAST	005022 saccharomyc
36	95.5	6.9	431	1	ARO4_AQUAE	067494 aquifex aeo
37	95.5	6.9	684	1	FLID_HELPY	06786 helicobacte
38	95.5	6.9	1013	1	SCA4_RICRH	09a181 rickettsia
39	95.5	6.9	1409	1	HAP1_HAEIN	P4596 haemophilus
40	95	6.9	382	1	VATC_MOUSE	002193 mus musculu
41	95	6.9	471	1	UDP5_ASTWE	091x97 astirgalus
42	95	6.9	513	1	LEU1_LACTA	002141 lactococcus
43	95	6.9	541	1	CH60_STRCO	09kxus streptococ
44	95	6.9	559	1	ILVD_PROMP	07vlt1 prochloroco
45	94.5	6.8	919	1	ATC1_RAT	064566 rattus norv

RESULT 1

AC	PLPA_PASHA	STANDARD;	PRT;	277 AA.
AC	008668; Q07363;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Outer membrane lipoprotein 1 precursor (PLP1).			
GN	PLPA.			
OS	Pasteurella haemolytica.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			
OC	Pasteurellaceae; Mannheimia.			
OX	NCBI_TaxID=75985;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Serotype A1;			
RC	MEDLINE=9332810; PubMed=8335249;			
RA	Murphy G.V., Whitworth L.C.;			
RT	"Analysis of tandem, multiple genes encoding 30-kDa membrane proteins in Pasteurella haemolytica A1."			
RL	Gene 129:107-111 (1993).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Serotype A1;			
RC	MEDLINE=94011378; PubMed=8406866;			
RA	Cooney B.J., Lo R.Y.C.;			
RT	"Three contiguous lipoprotein genes in Pasteurella haemolytica A1 type b."			
RL	Infect. Immun. 61:4682-4688 (1993).			
RM	-1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor (Probable).			
CC	-1- SIMILARITY: Belongs to the nlpA lipoprotein family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L11037; AAA25538.1; -			
DR	EMBL; M91072; AAA25541.1; -			
DR	EMBL; U16277; AAA25546.1; -			
DR	PIR; J00751; J00751.			
DR	InterPro; IPR004872; Lipoprotein_9.			
DR	InterPro; IPR004437; Prok_lipoprot_5.			
DR	InterPro; IPR004478; YaeC.			
DR	Pfam; PF01180; Lipoprotein_9; 1.			
DR	TIGRFAMS; TIGR00363; TIGR00363; 1.			
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.			
DR	Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	277	OUTER MEMBRANE LIPOPROTEIN 1.
FT	LIPID	20	20	N-palmitoyl cysteine (Probable)

FT LIPID 20 S-diacylglycerol cysteine (Probable).
FT CONFLICT 30 A -> R (IN REF. 2).
FT CONFLICT 156 L -> LIL (IN REF. 2).
FT CONFLICT 160 Q -> L (IN REF. 2).
FT CONFLICT 197 KALD -> NVS (IN REF. 2).
SQ SEQUENCE 277 AA; 25992 MW; 75387650306126 CRC64;
Query Match 73.2%; Score 1012.5; DB 1; Length 277;
Best Local Similarity 71.6%; Pred. No. 5,9e-61;
Matches 199; Conservative 32; Mismatches 44; Indels 3; Gaps 2;
QY 1 MNFGKINGICALASGIALAGCSNQ--SNEPAISKTAAOTIKYGVAPGEQVAEVAQGV 58
DB 1 MSFKKLIGV-ALVSAALATACKKEKKAESTAPAAQAPAKIKYGVSGSEHTVAERAAOI 59
QY 59 AKKKYLVTELVFENDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLIYGNFV 118
DB 60 AKKKYLVTELVFENDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLIYGNFV 119
QY 119 YPLAGYSTKIKTLNELKDGATTAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLDIV 178
DB 120 YPLAGYSKKYKVNSELAEAGVAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTVDII 179
QY 179 ENKKYVKEVDTSVAAARAIIDVDLAVNNNTYAGVGLASNGVVEKDSPPYNIIVA 238
DB 180 ENKKYVKEVDTSVAAARAIIDVDLAVNNNTYAGVGLASNGVVEKDSPPYNIIVA 239
QY 239 RADNDSKAIQDFKAVQCTDEVEAEAKQOFKQGVIKGW 276
DB 240 RADNDSKAIQDFKAVQCTDEVEAEAKQOFKQGVIKGW 277
RESULT 2
METQ_HABIN STANDARD; PRT; 273 AA.
ID METQ_HABIN 21 273
AC P31728;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable D-methionine-binding lipoprotein metQ precursor
DE (28 kDa outer membrane protein).
GN METQ OR H1620.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
ON [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype B;
RX MEDLINE=9110034; PubMed=1987077;
RA Chanyangam M., Smith A.L., Moseley S.L., Kuehn M., Jenny P.;
RT "Contribution of a 28-kilodalton membrane protein to the virulence of
RT Haemophilus influenzae.";
RL Infect. Immun. 59:600-608 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischnann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Greim C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512 (1995).
RN [3]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;

RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-428 (2000).
CC -1- FUNCTION: This protein is a component of a D-methionine
CC peptidase, a binding protein-dependent, ATP-driven transport
CC system (by similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59604; AAA24939.1; -;
CC EMBL; U32744; AAC22279.1; -;
CC PIR; B64082; B64082.
CC TIGR; H10620; -;
CC InterPro; IPR004872; Lipoprotein_9.
CC InterPro; IPR000437; Prok_lipoprot_9.
CC InterPro; IPR004478; YaeC.
CC Pfam; PF01180; Lipoprotein_9; 1.
CC TIGRFAMs; TIGR00363; TIGR00363; 1.
CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC Transport; Amino-acid transport; Outer membrane; Lipoprotein; Signal;
CC Complete proteome; Palmitate.
CC SIGNAL 1 20 PROBABLE.
CC CHAIN 21 273 PROBABLE D-METHIONINE-BINDING LIPOPROTEIN
FT METQ.
FT LIPID 21 21 N-palmitoyl cysteine (Probable).
FT CONFLICT 249 249 S-diacylglycerol cysteine (Probable).
SQ SEQUENCE 273 AA; 23928 MW; 429938A8AC7DD7D7 CRC64;
Query Match 72.8%; Score 1006.5; DB 1; Length 273;
Best Local Similarity 71.4%; Pred. No. 1,5e-60;
Matches 197; Conservative 28; Mismatches 48; Indels 3; Gaps 1;
QY 1 MNFGKINGICALASGIALAGCSNQSNNEPAISKTAAOTIKYGVAPGEQVAEVAQGVAK 60
DB 1 MKLKQPLATRAIASALVLTGCKEDKKPEAA--AAPFKIKYGVSGSEHTVAERAAQVAK 57
QY 61 EKINLVTELVFENDYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNLIYGNFVYP 120
DB 58 EKVGLDVQFVEFNDYALPNEAVSKGELDANAMQHKPYLEKDSQKGLNLIYGNFVYP 117
QY 121 LAGSTKIKTLNELKDGATTAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLDIVEN 180
DB 118 LAGSTKIKTLNELKDGATTAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLDIVEN 177
QY 181 PKKLVKEVDTSVAAARAIIDVDLAVNNNTYAGVGLTASENGVVEKDSPPYNIIVA 240
DB 178 PKKLVKEVDTSVAAARAIIDVDLAVNNNTYAGVGLTASENGVVEKDSPPYNIIVSRT 237
QY 241 DNKDSKAIQDFKAVQCTDEVEAEAKQOFKQGVIKGW 276
DB 238 DNKDSKAIQDFKAVQCTDEVEAEAKQOFKQGVIKGW 273
RESULT 3
METQ_PASMTU STANDARD; PRT; 276 AA.
ID METQ_PASMTU 21 276
AC Q9CK95;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable D-methionine-binding lipoprotein metQ precursor
DE (Outer membrane lipoprotein 1).

GN METO OR PLPA OR PM1730.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 perase, a binding protein-dependent, ATP-driven transport
 system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 anchor (Probable).
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB06210; AK03814.1; -
 CC InterPro: IPR004872; Lipoprotein_9.
 CC InterPro: IPR004872; Prok_lipoprot_9.
 CC InterPro: IPR004478; YaeC.
 CC Pfam: PF03180; Lipoprotein_9; 1.
 CC TIGRPFAMS: TIGR00363; TIGR00363; 1.
 CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Transport: Amino-acid transport; Outer membrane; Lipoprotein; Signal;
 CC Complete proteome; Palmitate.
 CC FT SIGNAL 1 20 PROBABLE.
 CC FT CHAIN 1 276 PROBABLE D-METHIONINE-BINDING LIPOPROTEIN
 CC FT LIPID 21 21 METO.
 CC FT LIPID 21 21 N-palmitoyl cysteine (Probable).
 CC FT LIPID 21 21 S-diacylglycerol cysteine (Probable).
 CC SQ SEQUENCE 276 AA; 30232 MW; 7AFD62A687D624 CRC64;
 Query Match 64.2%; Score 888; DB 1; Length 276;
 Best Local Similarity 60.2%; Pred. No. 1.3e-52;
 Matches 168; Conservative 55; Mismatches 50; Indels 6; Gaps 3;
 QY 1 MNFGINGICALASGIALAGCSNOSNEBPAIS---KTAAGITIKYVWAGPEQVAEVAAGQ 57
 DB 1 MTLTLPGLATLTVAVALLAGC--KODKPAALAAPOEPAPARKLTGVWMTGAAQAQTEVAAK 58
 QY 58 VAKKXNLTVELVEENDYAMPNSAVSKGLDANMOMHPRYLEKDSQEKGLNLTIVGNTF 117
 DB 59 IAKKXNIDVLTVEFTETQPNDAITKGLDANMFMHPRVMDKVEGRGY-KLAIIVGNTF 117
 QY 118 VYPLAGVSTKTKTINELDQATIVPNPNSUARALLILEQGLIKKDNNTNLFSTLIDI 177
 DB 118 VPLIAYSKTKIKNVELODQATVAVPNPNSUARALLILEQGLIKKDNNTNLFSTLIDI 177
 QY 178 VEPKRLVKEVDVSAARALDIDVLAIVNNVYAGQVGLTASNGVFEVDEDSRYNNIIV 237
 DB 178 IENPKNLDKEVEGSLPRMLDDVDFAIINNVAVQGLTEKGIIVEDDSRYNNIIV 237
 QY 238 ARADNKSQALQDPYKAVQTEVEAEAKQKQKGYIKGM 276
 DB 238 SREDNKDNEAIKDFYKAFQTEVEVYQEALEKHPQGGVYKGM 276

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metc precursor.
 GN METO OR STY0272 OR T0248.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltham T., Hamlin N., Hargreaves A., Hen T.T., Holroyd S., Jorgensen K.,
 RA Krogan A., Larsen T.S., Leather S., Mouton S., O'Garra P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skellon J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=TY2 / ATCC 700931;
 CC MEDLINE=22531367; PubMed=12644504;
 CC Deng W., Lion S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 CC Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
 CC "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 CC and CT18.";
 CC J. Bacteriol. 185:2330-2337(2003).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 perase, a binding protein-dependent, ATP-driven transport
 system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MISCELLANEOUS: The metNIO system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL627266; CAD08705.1; -
 CC EMBL: AB016834; AAC67977.1; -
 CC InterPro: IPR004872; Lipoprotein_9.
 CC InterPro: IPR004872; Prok_lipoprot_9.
 CC InterPro: IPR004478; YaeC.
 CC Pfam: PF03180; Lipoprotein_9; 1.
 CC TIGRPFAMS: TIGR00363; TIGR00363; 1.
 CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Transport: Amino-acid transport; Membrane; Lipoprotein; Signal;
 CC Complete proteome; Palmitate.
 CC FT SIGNAL 1 22 POTENTIAL.
 CC FT CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METO.
 CC FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 CC FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 CC SQ SEQUENCE 271 AA; 29465 MW; 49B62C4CF96D9613 CRC64;
 Query Match 57.4%; Score 793.5; DB 1; Length 271;
 Best Local Similarity 56.5%; Pred. No. 2.6e-46;
 Matches 156; Conservative 44; Mismatches 65; Indels 11; Gaps 3;
 QY 3 FGKINGICALASGIALAGCSNOSNEBPAISKTAAGITIKYVWAGPEQVAEVAAGVAKEX 62
 DB 5 FKTPAVGALLIGSIALAGCGDDEKDP-----NHIKGVIVGAEQVAVAEVAGVAKEX 56

```

QY 63 YNLFEVLEFNDYAMPNSAVSGKGLDANAMQHKPELEKDSQEKGLNNLVIGNTFYVPLA 122
DB 57 YGLDVELVTNDYVLPNPAALSKGIDIDANAFQHKPYLDQIKDNGY-KLVSVGKTFYVPLA 115
QY 123 GYSTKIKTLNELKOGATIAVNDPSNLARALILKFGKGLKDKDNTNLTSTLDIVENPK 182
DB 116 GYSKIKSIDELDKGSGQVAVPNDPTNLGRSLLLQKVGILLKDKGSLPTSLDIYENPK 175
QY 183 KLVIKEVDTSVAARAIID--VDLAVVNNNYAGQVGLTASNGVFEEDKDSFYVNIIVARA 240
DB 176 NLKIVLEAPQLFRSLDDAQIALAVINTTYASQIGLTPAKDGI FVEDKDSFYVNIIVTRE 235
QY 241 DNKDSKAIODFYKAYQTDVEAEAKKQPKDGVKGM 276
DB 236 DNKDAENVKKEFYQAYQSDVEYEAANKVFNCGAVKGM 271

RESULT 5
METO_SALTY
ID METO_SALTY STANDARD; PRT; 271 AA.
AC 082RN1:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB D-methionine-binding lipoprotein meto precursor.
GN METO OR STM0245.
OS Salmoneella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmoneella.
OX NCBI_TaxID=602;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=L72 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Rader E., Sun H., Florea L., Miller W., Stuckert T., Nian M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmoneella enterica serovar Typhimurium
RT L72."
RL Nature 413:852-856 (2001).
CC -!- FUNCTION: This protein is a component of a D-methionine
CC permease, a binding protein-dependent, ATP-driven transport
CC system (by similarity).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -!- MISCELLANEOUS: The metNIQ system is also able to transport
CC the toxic methionine analog alpha-methyl-methionine (by
CC similarity).
CC -!- SIMILARITY: Belongs to the nlpA lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AF008706; AAL19208.1; -
DB StyGene; SG7777; metQ.
DB InterPro; IPR004872; Lipoprotein_9.
DB InterPro; IPR004437; Prok_Lipoprot_S.
DB InterPro; IPR004478; YaeC.
DB Pfam; PF01180; Lipoprotein_9; 1.
DB TIGRfams; TIGR00363; TIGR00363; 1.
DB PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DB Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
DB Complete proteome; Palmitate.
FT SIGNAL 1 22
FT CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METO.

```

```

FT LIPID 23 23 N-palmitoyl cysteine (Potential).
FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
SQ SEQUENCE 271 AA; 29435 MW; 3F6D41E14FBD707 CRC64;
Query Match. 57.3%; Score 792.5; DB 1; Length 271;
Beet Local Similarity 56.5%; Pred. No. 3e-46; Indels 11; Gaps 3;
Matches 156; Conservative 43; Mismatches 66;

QY 3 FGKINGICALASGIALAGCSNQSNEPAISKTAQTIKGVNAGSEQAAVEAQAQVAREK 62
DB 5 FKTPAAVAGLIGSLALAGCGQDEKDP-----NHIKGVIVGAEOQVAEVAQKVAEK 56
QY YNLFEVLEFNDYAMPNSAVSGKGLDANAMQHKPELEKDSQEKGLNNLVIGNTFYVPLA 122
DB 57 YGLDVELVTNDYVLPNPAALSKGIDIDANAFQHKPYLDQIKDNGY-KLVSVGKTFYVPLA 115
QY 123 GYSTKIKTLNELKOGATIAVNDPSNLARALILKFGKGLKDKDNTNLTSTLDIVENPK 182
DB 116 GYSKIKSIDELDKGSGQVAVPNDPTNLGRSLLLQKVGILLKDKGSLPTSLDIYENPK 175
QY 183 KLVIKEVDTSVAARAIID--VDLAVVNNNYAGQVGLTASNGVFEEDKDSFYVNIIVARA 240
DB 176 NLKIVLEAPQLFRSLDDAQIALAVINTTYASQIGLTPAKDGI FVEDKDSFYVNIIVTRE 235
QY 241 DNKDSKAIODFYKAYQTDVEAEAKKQPKDGVKGM 276
DB 236 DNKDAENVKKEFYQAYQSDVEYEAANKVFNCGAVKGM 271

RESULT 6
PLPC_PASHA
ID PLPC_PASHA STANDARD; PRT; 263 AA.
AC Q06870; Q07365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Outer membrane lipoprotein 3 precursor (Plp3).
GN PlpC.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=93328110; PubMed=8335249;
RA Murphy G.L., Whitworth L.C.;
RT "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins
RT in Pasteurella haemolytica A1."
RL Gene 129:107-111 (1993).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=84011378; PubMed=8406866;
RA Cooney B.J., Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
RT which are homologous to a lipoprotein gene in Haemophilus influenzae
RT type b."
RL Infect. Immun. 61:4682-4688 (1993).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC -!- SIMILARITY: Belongs to the nlpA lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; L11037; AAA25540.1; -
DB EMBL; L16627; AAA25548.1; ALT_INIT.

```

DR PIR; JN0753; JN0753.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_LipProt_S.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR TIGRfams; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.
 FT SIGNAL 1 19
 FT CHAIN 20 263
 FT LIPID 20 20
 FT LIPID 20 20
 FT LIPID 20 20
 FT CONFLICT 103 103
 FT CONFLICT 200 208
 SQ SEQUENCE 263 AA; 29093 MW; 68FDD8460ED579CB CRC64;

Query Match 57.2%; Score 791.5; DB 1; Length 263;
 Best Local Similarity 56.9%; Pred. No. 3.4e-46;
 Matches 157; Conservative 49; Mismatches 57; Indels 13; Gaps 3;

QY 1 MNFGKINGICALASGIALAGCSNSENPAISKTAQTITKGVMAQPEQVAEVAAGQYAK 60
 DB 1 MKIMKAGVALIFS-LFTACNDK-----AEKLVGVISGPEKMEVAKIAK 48
 QY 61 EKNTLVELVEFNDAVMPNSAVSKGELDANAMQHKPYLEKDSQKGLNNLVVNGTFYYP 120
 DB 49 EKNREDVELVFTDVAATPNAALDKGLDLNAFQHKPYLDNQIQEKGY-KLVAVGNTFYYP 107
 QY 121 LAGYKTKITNELKDGATIAVPNDPSNLARALLILEKQGLIKLKDNTLPSFTLIDVEN 180
 DB 108 IAYSKIKSLAELEKDGDTIAVPNDPTNLARALLILEKQGLIKLADAGLKATSDIEN 167
 QY 181 PKLVKEVDTSVARAIDVDLAANNVYAGOVGLTASENFVEEDKSPVNIIVARA 240
 DB 168 PKLVIGELAEPLRLTDDVAFSIINTTYAQNGITPKDSIFVEDKSPVNIIVARE 227
 QY 241 DNKSKAIDPFYKAYOTDEVEAEAKKQFDGVYIKGM 276
 DB 228 NNQSEAVDLVKAQYOTDEVYKANEFEKGMIKGM 263

RESULT 7

METQ_ECO57
 ID METQ_ECO57 STANDARD; PRT; 271 AA.

AC Q8XSV9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR Z0209 OR ECS0199.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN RP
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDJ933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoustis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533 (2001).
 RN RP
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11259796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsuo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
 RA Tida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,

RA Kulara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MISCELLANEOUS: The metQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the nipa lipoprotein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE005195; AAG54499.1; -
 DR EMBL; AP002550; BAB33622.1; -
 DR PIR; G85504; G85504.
 DR PIR; G80653; G80653.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_LipProt_S.
 DR InterPro: IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRfams; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1 22
 FT CHAIN 23 271
 FT LIPID 23 23
 FT LIPID 23 23
 FT LIPID 23 23
 SQ SEQUENCE 271 AA; 29416 MW; 39D7DALS5CD2BFB CRC64;

Query Match 57.2%; Score 791.5; DB 1; Length 271;
 Best Local Similarity 55.8%; Pred. No. 3.5e-46;
 Matches 154; Conservative 44; Mismatches 67; Indels 11; Gaps 3;

QY 3 FGKINGICALLASGIALAGCSNSENPAISKTAQTITKGVMAQPEQVAEVAAGQYAK 62
 DB 5 FKTFAVAGALIGSLAVGGQDEKDP-----NHKIGVIVGAQGVAEVAQYAKK 56
 QY 63 YNLTVELVEFNDAVMPNSAVSKGELDANAMQHKPYLEKDSQKGLNNLVVNGTFYYP 122
 DB 57 YGLDVELVFTDVAATPNAALDKGLDLNAFQHKPYLDNQIQEKGY-KLVAVGNTFYYP 115
 QY 123 GYSTKTKITNELKDGATIAVPNDPSNLARALLILEKQGLIKLKDNTLPSFTLIDVEN 182
 DB 116 GYSKTKISLDELDDGQVAVPNDPTNLGRSLILLLQVGLIKLKDGVGLPLTVLDVVEN 175
 QY 183 KLVKEVDTSVARAIDVDLAANNVYAGOVGLTASENFVEEDKSPVNIIVARA 240
 DB 176 NKLIVLEAPQPLRSDDQIALAVNTTYAQGLTPKMGIFVEDKSPVNIIVARE 235
 QY 241 DNKSKAIDPFYKAYOTDEVEAEAKKQFDGVYIKGM 276
 DB 236 DNKDAENVKFFVQAYQSDDEVYBAANKVFNGAVKGM 271

RESULT 8

METQ_ECOLI
 ID METQ_ECOLI STANDARD; PRT; 271 AA.

AC F28635;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.

GN METO OR B0197.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RL Miyamoto K.;
 RM Submitted (Apr-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RL Miyamoto K.;
 RM Submitted (Apr-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / M3110;
 RL Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada B., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of the
 4.0 - 6.0 min (189,987 - 281,416bp) region.";
 RN Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
 RA Laskari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
 RA Davis R.W.;
 RN Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 127-271 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93094132; PubMed=1459951;
 RA Gervais F.G., Drapeau G.R.;
 RT "Identification, cloning, and characterization of rcsF, a new
 regulator gene for exopolysaccharide synthesis that suppresses the
 division mutation fts284 in Escherichia coli K-12.";
 RN J. Bacteriol. 174:8016-8022(1992).
 RN [6]
 RP FUNCTION.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=22159550; PubMed=12169620;
 RA Gal U., Szeverik A., Schnell R., Kalman M.;
 RT "The metC D-methionine transporter locus of Escherichia coli is an ABC
 transporter gene cluster.";
 RN J. Bacteriol. 184:4930-4932(2002).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 permease, a binding protein-dependent, ATP-driven transport
 system.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Probable).
 CC -1- MISCELLANEOUS: The metNIG system is also to be able to transport
 the toxic methionine analog alpha-methyl-methionine.
 CC -1- SIMILARITY: Belongs to the nlpA lipoprotein family.
 CC -1- CAUTION: REP. 5 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITION 217
 ONWARD DUE TO A FRAMESHIFT.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; D15061; BAA03657.1; -

DR EMBL; AE000129; AAC73308.1; -
 DR EMBL; D83536; BAA77874.1; -
 DR EMBL; U70214; AAB08625.1; -
 DR EMBL; L04474; AAA24507.1; ALT_FRAME.
 DR PIR; E64744; E64744.
 DR SWISS-2DPAGE; P28635; COLI.
 DR Ecogen; EGI504; metQ.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_Lipoprot_S.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF01180; Lipoprotein_9; 1.
 DR TIGRfam; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 DR Transprot; amino-acid transport; Membrane; Lipoprotein; signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1 22
 FT CHAIN 23 271
 FT LIPID 23 23
 FT LIPID 23 23
 FT CONFLICT 200 200
 FT SEQUENCE 271 AA; 29431 MW; B50CBCEB5CD2BF7 CRC64;
 Query Match 56.6%; Score 782.5; DB 1; Length 271;
 Best Local Similarity 55.1%; Pred. No. 1.4e-45;
 Matches 152; Conservative 46; Mismatches 67; Indels 11; Gaps 3;
 QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQAVAEVAGQVAREK 62
 DB 5 FKTFPAAGALIGSLALVGGQGEKDP-----NHKGVGVGAEGQVAVAGQVARDK 56
 QY 63 YNTLVEFENDPAMNSAVSKGELDANMOKRPLEHDSQSKGANNVYQNTFVYLA 122
 DB 57 YGDLVELVFNDDVLENEALISKGDIDANAFQKPYLDQDLKRGY-KVAAGNTFVYLA 115
 QY 123 GYSTKIKTNELKDGATIAVPDPNSMLAPALILKKGILKDKNTNLFSTLIDVENPK 182
 DB 116 GYSKIKRISDELDDGQVAVPDPNTLGRSLLLQKVGILKDKGVGLPFLVDVENPK 175
 QY 183 KLYIKERVDTSVARRADD--VDLAVNNNYACQVGLTASENVFPEEDSPVNIIVARA 240
 DB 176 NLKIVLEAPQLFRSLDDQALAVANTVYASQIGLTPKQGIFFEDKESPVNLTVRE 235
 QY 241 DNKSKAIDPFVAYQTDVEAEAKKQPFQGYIKGM 276
 DB 236 DNKDAENVKKFVQAYQSDVEYEAANKVFNGAVKGM 271
 RESULT 9
 PLPB_PASHA STANDARD; PRT; 276 AA.
 ID PLPB_PASHA STANDARD; PRT; 276 AA.
 AC Q08869; Q07364;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Outer membrane lipoprotein 2 precursor (PLP2).
 GN PLPB.
 OS Pasteurella haemolytica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=75985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype A1;
 RX MEDLINE=93328110; PubMed=8335249;
 RA Murphy G.L., Whitworth D.C.;
 RT "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins
 in Pasteurella haemolytica A1.";
 RN Gene 129:107-111(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serotype A1;
 RX MEDLINE=94011378; PubMed=8406866;
 RA Cooney B.J., Lo R.Y.C.;

RT "Three contiguous lipoprotein genes in *Pasteurella haemolytica* A1
 RT type b.",
 RT are homologous to a lipoprotein gene in *Haemophilus influenzae*
 RL Infect. Immun. 61:4692-4698(1993).
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor (Probable).
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L11037; AAA25539.1; -
 DR EMBL: L16627; AAA25547.1; -
 DR PIR: J00752; J00752.
 DR InterPro: IPR004872; Lipoprotein 9.
 DR InterPro: IPR000437; Prok_lipoprot_5.
 DR InterPro: IPR004478; YaeC
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRfam: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Outer membrane; Lipoprotein; signal; Multigene family; Palmitate.
 KW SIGNAL. 1
 FT CHAIN 19
 FT LIPID 20 276 N-palmitoyl cysteine (Probable).
 FT LIPID 20 20 S-diacylglycerol cysteine (Probable).
 FT CONFLICT 268 270 FNG -> LVTH (IN REF. 2).
 FT SEQUENCE 276 AA; 30157 MW; DFBP443ADAD4D950 CRC64;
 SQ
 Query Match 55.9%; Score 773; DB 1; Length 276;
 Best Local Similarity 54.0%; Pred. No. 6,3e-45;
 Matches 150; Conservative 54; Mismatches 70; Indels 4; Gaps 4;
 QY 1 MNFGKINGICAGIALAGCSNQ-SNEPAISKTAQT-INVGVAGGEQVAEYAGV 58
 DB 1 MNFGKINGICAGIALAGCSNQ-SNEPAISKTAQT-INVGVAGGEQVAEYAGV 58
 QY 1 MNFGKINGICAGIALAGCSNQ-SNEPAISKTAQT-INVGVAGGEQVAEYAGV 58
 DB 1 MNFGKINGICAGIALAGCSNQ-SNEPAISKTAQT-INVGVAGGEQVAEYAGV 58
 QY 59 AAEKXVLTVEIVFENDYAMPNSAVSGEIDANAMOHKPYLEDSOEKGNLVIGNTFV 118
 DB 60 AAEKXVLTVEIVFENDYAMPNSAVSGEIDANAMOHKPYLEDSOEKGNLVIGNTFV 118
 QY 119 YPLAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTNLFSTLDIV 178
 DB 119 YPLAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTNLFSTLDIV 178
 QY 119 WPIAAVSKIKINSELDKGTAVAPNNAINTARALLLGAHGLIKLKDKKVFATENDII 178
 DB 119 WPIAAVSKIKINSELDKGTAVAPNNAINTARALLLGAHGLIKLKDKKVFATENDII 178
 QY 179 ENPKKLVIEVNTSVAAARIDVDVLA VNNNTNAGVGLTASENGVFVEKDSPPYNIYA 238
 DB 179 ENPKKLVIEVNTSVAAARIDVDVLA VNNNTNAGVGLTASENGVFVEKDSPPYNIYA 238
 QY 239 RADNKSKAIDPFVAKYQTFEVEAEAKQFVKGVIKGW 276
 DB 239 RADNKSKAIDPFVAKYQTFEVEAEAKQFVKGVIKGW 276
 QY 239 REDNNDPRLQITVVSFOTIEVFGELAKLFGNGVYKGW 276
 DB 239 REDNNDPRLQITVVSFOTIEVFGELAKLFGNGVYKGW 276
 RESULT 10
 METQ YERP
 ID METQ YERP STANDARD; PRT; 271 AA.
 AC 082H40;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR YPO1071 OR Y3106.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OC NCBI_TaxId=632;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CO-92 / Bivovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parthill U., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Meule S., Oyston P.C.F., Quail M.A., Rutherford G.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,
 RT Genome sequence of *Yersinia pestis*, the causative agent of plague.;
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KIMS / Bivovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RA "Genome sequence of *Yersinia pestis* KIM.";
 RL Bacteriol. 184:4601-4611(2002).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (by similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MISCELLANEOUS: The metIQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine (by
 CC similarity).
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ141416; CAC99914.1; -
 DR EMBL: AE013911; AAM8656.1; -
 DR PIR: AG0131; AG0131.
 DR InterPro: IPR004872; Lipoprotein 9.
 DR InterPro: IPR000437; Prok_lipoprot_5.
 DR InterPro: IPR004478; YaeC
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRfam: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Membrane; Lipoprotein; signal;
 KW Complete proteome; Palmitate.
 KW SIGNAL. 1
 FT CHAIN 22
 FT LIPID 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
 FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 FT SEQUENCE 271 AA; 29376 MW; S214CEB05117FC8 CRC64;
 SQ
 Query Match 54.8%; Score 757.5; DB 1; Length 271;
 Best Local Similarity 53.2%; Pred. No. 6,7e-44;
 Matches 148; Conservative 46; Mismatches 71; Indels 11; Gaps 3;
 QY 1 MNFGKINGICAGIALAGCSNQ-SNEPAISKTAQT-INVGVAGGEQVAEYAGV 60
 DB 3 LKFSIAISLIGTLVVGQPTKAP-----NHIVGVIVGEQVAEYAGV 54
 QY 61 EKXNLTVEIVFENDYAMPNSAVSGEIDANAMOHKPYLEDSOEKGNLVIGNTFV 120
 DB 55 EKXVLTVEIVFENDYVLPNEALSIGDIDINAFQHKPYLDQIKDRG-KLVVSGNSFV 113
 QY 121 LAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTNLFSTLDIV 180
 DB 114 LAGYSTIKITLNEKDGATIAVPDPSNLARALILEKQGLIKLKDNTNLFSTLDIV 173

106 0011 22 11.07.07 2007

QY 181 PKTLVKEVDTSVAARAID--VDLAVNNNNAAGVGLTASNGVFVEDKSDSPYNNIYA 238
DB 174 PKNLKVELEAPLPSPDLDDQIALAIINTTAAAGTGLTPADGLFVEDKSPYNNLIVA 233
QY 239 RADNDSKAIQDFVKAQYQTEVEAEAKQFKQGVIKGM 276
DB 234 REDNKAENKFKVQAYQSDVEYDAANKAFNGAVKGM 271

RESULT 11
METH_VIBCH STANDARD; PRT; 269 AA.
ID METH_VIBCH
AC 09KTD7;
DT 28-FEB-2003 (Rel. 41, Last created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable D-methionine-binding lipoprotein metQ precursor.
GN METQ OR VC0305.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Bisen J.A., Nelson M.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.E., Hickey E.K., Peterson J.D., Umagam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Esmolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Ueberack T., Fleischmann R.D., Niemann M.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.D., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: This protein is a component of a D-methionine
CC permease, a binding protein-dependent, ATP-driven transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004174; AAF94067.1; ALT_INIT.
DR TIGR; VCO905; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR004478; Yaec.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRFAMs; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
KW Complete proteome; Palmitate.
FT SIGNAL 1 22
FT CHAIN 1 269
FT LIPID 23 23
FT LIPID 23 23
SQ SEQUENCE 269 AA; 29067 MW; 39E2570BE3F184D6 CRC64;
Query Match 54.3%; Score 751.5; DB 1; Length 269;
Best Local Similarity 54.1%; Pred. No. 1,7e-43;
Matches 145; Conservative 51; Mismatches 63; Indels 9; Gaps 2;
QY 9 ICALASGIALAGSGNSGNEPALSKTAQTIKVWAGPQAVAEVAGVAKKYNLTVE 68

DB 11 IAAASLTLLAGCKEKAVD-----NNKVKIGWAGAGVAAQVAAKAKKYNLDVE 62
QY 69 LVEFNDAENSAVSKKELDANAMQHPYLEKSOEGLNNLYVNGTFPYPLAGSTKI 126
DB 63 LVTFPTDVTNPAALDDSIDANAFQHPYLDKQIADRGY-KLAVGTFVYPIAGISKQI 121
QY 129 KTLNELKQATIAVNPDPNSMLARALLERKQGIKDKONTNLTSTLDIYENPKVIXE 168
DB 122 KSVDELQDGAIRIVNPDPNTLGRSLILLLQGGILKLRBDVGLLATVADYENPKVILE 161
QY 189 VDTSVAAARAIDVDLAVNNNNAAGVGLTASNGVFVEDKSDSPYNNIYARANKSKAI 248
DB 182 LDAQQLPRSLDDVALSTINTTAAAGTGLTPEKQGFVENKESPYNNLIVAREANVAENV 241
QY 249 QDFVKAQYQTEVEAEAKQFKQGVIKGM 276
DB 242 QNFKKAQYQTEVEAEAEITFGGAVKGM 269

RESULT 12
NLPA_ECOLI STANDARD; PRT; 272 AA.
ID NLPA_ECOLI
AC P04876;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lipoprotein-28 precursor.
GN NLPA OR B3661.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86111928; PubMed=3003106;
RA Yu F., Inouye S., Inouye M.;
RT "Lipoprotein-28, a cytoplasmic membrane lipoprotein from Escherichia
RT coli. Cloning, DNA sequence, and expression of its gene";
RL J. Biol. Chem. 261:2284-2288(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KL2 / MG1655;
RX MEDLINE=93315143; PubMed=7686882;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
RT genome: organizational symmetry around the origin of replication";
RL Genomics 16:551-561(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
CC anchor.
CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M12163; AAA24080.1; -
DR EMBL; L10328; AAA62013.1; -
DR EMBL; AE000443; AAC76684.1; -
DR PIR; A26286; LPEC28.
DR SWISS-2DPAGE; P04846; COLI.
DR Ecocore; EG10657; nlpa.
DR InterPro; IPR004672; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR004478; Yaec.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRFAMs; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Inner membrane; Lipoprotein; Signal; Complete proteome; Palmitate.
KM

FT SIGNAL 1 23
 FT CHAIN 24 272 LIPOPROTEIN-28.
 FT LIPID 24 24 N-palmitoyl cysteine.
 FT LIPID 24 24 S-diacylglycerol cysteine.
 SQ SEQUENCE 272 AA; E3C249E753AB1B33 CRC64;

Query Match 50.0%; Score 691.5; DB 1; Length 272;
 Best Local Similarity 50.6%; Pred. No. 1,8e-39;
 Matches 137; Conservative 53; Mismatches 70; Indels 11; Gaps 3;

QY 8 GICALASGIALAGCSNQSNEPAISKTAAQTIKVGWAGPEQAAVEAGVAKKNTLV 67
 DB 11 GAALLAGILLAGCQSSD-----AKHIKVGIVNGAEQDVAVAKKAKKYGGLDV 62
 QY 68 ELVEENDYAMPNSAVSKGELDANAMQHPYLEKDSQELNNLVYVNTFYVPLAGYSTK 127
 DB 63 ELVGFSGSLPNDATNHELDANVFQHPLEQDNQAHGY-KLVAVGNITVFPVAGYSKK 121
 QY 128 IKTLNELKDGATIAVPNDPSNLARALITLKEKGLIKLNDNTNLFSTLIDIVENPKVLV 187
 DB 122 IKTVAQIKEGATVAILPNDPTNLGRALLLQKEKILTLKEKGLLPTALDITDNPRLQIM 181
 QY 188 EVDTSVAAALD--VDLAVNNNYAGVGLTASNGVFEVFKDSPYVNIIVARADNDS 245
 DB 182 ELBGAQLPRLVDPKVDVALISTTYIQGTGLSPVHDSVFIEDKNSPVNIIIVAREDNKA 241
 QY 246 KAIDPFVAVOTDEVEAEAKKQFQDGVIKGW 276
 DB 242 ENVKEFLOSYSPEVAKAAETIFNGAVPGW 272

RESULT 13
 NLPB_ECO57 STANDARD; PRT; 272 AA.
 ID NLPB_ECO57
 AC 08XC50:
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lipoprotein-28 precursor.
 GN NLPB OR Z5147 OR ECS4595.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=83334;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=1120651;
 RA Perna N.T., Plunkett G., Burland V., Mau B., Glasner J.D.,
 RA Rose D.U., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller D.,
 RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamowski K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT Genome sequence of enterohemorrhagic Escherichia coli O157:H7.;
 RT Nature 409:529-533(2001).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shingawa H.,
 RT Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RL -1- SUBCELLULAR LOCATION: Attached to the inner membrane by a lipid
 anchor (By similarity).
 CC -1- SIMILARITY: Belongs to the nlpA lipoprotein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AE005597; AAG58856.1; -
 DR EMBL; AP002566; BAB38018.1; -
 DR PIR; C91203; C91203.
 DR PIR; D86049; D86049.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_Lipoprot_9.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRPFams; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KM Inner membrane; Lipoprotein; Signal; Complete proteome; Palmitate.
 FT SIGNAL 1 23
 FT CHAIN 24 272 LIPOPROTEIN-28 (By similarity).
 FT LIPID 24 24 N-palmitoyl cysteine.
 FT LIPID 24 24 S-diacylglycerol cysteine.
 SQ SEQUENCE 272 AA; A52B4C8576B443F5 CRC64;

Query Match 49.6%; Score 686.5; DB 1; Length 272;
 Best Local Similarity 49.8%; Pred. No. 3.8e-39;
 Matches 135; Conservative 54; Mismatches 71; Indels 11; Gaps 3;

QY 8 GICALASGIALAGCSNQSNEPAISKTAAQTIKVGWAGPEQAAVEAGVAKKNTLV 67
 DB 11 GAALLAGILLAGCQSSD-----EKHIKVGIVNGAEQDVAVAKKAKKYGGLDV 62
 QY 68 ELVEENDYAMPNSAVSKGELDANAMQHPYLEKDSQELNNLVYVNTFYVPLAGYSTK 127
 DB 63 ELVGFSGSLPNDATNHELDANVFQHPLEQDNQAHGY-KLVAVGNITVFPVAGYSKK 121
 QY 128 IKTLNELKDGATIAVPNDPSNLARALITLKEKGLIKLNDNTNLFSTLIDIVENPKVLV 187
 DB 122 IKTVAQIKEGATVAILPNDPTNLGRALLLQKEKILTLKEKGLLPTALDITDNPRLQIM 181
 QY 188 EVDTSVAAALD--VDLAVNNNYAGVGLTASNGVFEVFKDSPYVNIIVARADNDS 245
 DB 182 ELBGAQLPRLVDPKVDVALISTTYIQGTGLSPVHDSVFIEDKNSPVNIIIVAREDNKA 241
 QY 246 KAIDPFVAVOTDEVEAEAKKQFQDGVIKGW 276
 DB 242 ENVKEFLOSYSPEVAKAAETIFNGAVPGW 272

RESULT 14
 TP32_TREPA STANDARD; PRT; 268 AA.
 ID TP32_TREPA
 AC 007950:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Membrane lipoprotein Tpn32 precursor (29 kDa protein).
 GN TP32 OR TP0821.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
 NCBI_TaxID=160;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RA Porcella S.F., Radolf J.D., Norgard M.V.;
 RT "Treponema pallidum 29k protein is homologous to a lipoprotein
 RT present in Pasteurella hemolytica and in Hemophilus influenzae
 RT type D.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

CC Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC CC

[31]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Kechum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salberg S., Petersen J.,
 RA Kralak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 RA McDonald L., Atciach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 RT spirochete."
 RL Science 281:375-388(1998).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U97358; AAB63362.1; -
 DR EMBL: U93844; AAB61267.1; -
 DR EMBL: AF001253; AAC65789.1; -
 DR PIR: G71275; G71275.
 DR TIGR: TP0821; -
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_lipoprot_9.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Membrane; Lipoprotein; Signal; Complete proteome; Palmitate.
 FT SIGNAL 1 23
 FT CHAIN 1 23 BY SIMILARITY.
 FT LIPID 24 24 MEMBRANE LIPOPROTEIN TP032.
 FT LIPID 24 24 N-palmitoyl cysteine (Potential).
 FT LIPID 24 24 S-diacylglycerol cysteine (Potential).
 SQ SEQUENCE 268 AA; 29081 MW; AB21F8A07A88F691 CRC64;
 Query Match 30.1%; Score 416.5; DB 1; Length 268;
 Best Local Similarity 34.1%; Pred. No. 4,4e-21;
 Matches 92; Conservative 61; Mismatches 100; Indels 17; Gaps 4;
 QY 4 GKNICGALASGIALAGCSNOSNEPAISKTAOTIKVGVMGPEQAAVEAVGQVAKERY 63
 DB 12 GKL-----IALGVVACTQVDE-----IVGVVSEPHARLLIAKEVYK-KQ 55
 QY 64 NLTVLEVENDYAMPSAVSKKELDANAMQHKPYLEKDSQEKLNINLVIVGTFYPIAG 123
 DB 56 HIELRLVEFTNYVALNEAVMRGDIIMNFQHPHMQCFQOEHN-GDLVSGVGHVEPLAL 114
 QY 124 YSTKLTLELKGATIAVNDPSNLARLILKEQGLIKKDNNTLFTSLDIYENPK 183
 DB 115 YSRTRHVDPPAGAVIALPNDSSNEARLRLLLEAAGFIRKAGSGLFATVEVDQOVEN 174
 QY 184 LVIKEDYTSVAAARAIIDVDLAVNNNYAQQVGLTASENGVEVDKSPYVNIIVAPADNK 243
 DB 175 VVLQEVESALLPRVDPQVGVAVINGVYALMAGLSARBDGLAVEPDASVAVLVYKRGNE 234
 QY 244 DSKAIQDPYKAYQTDVEAEAKQKRDGVI 273
 DB 235 ADARVQAVIRALCGGRVIRTYLKERKGEV 264
 RESULT 15
 YHCU YACU STANDARD; PRT; 263 AA.
 AC PS4534;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical lipoprotein yhcJ precursor.
 GN YHCU OR ESU09110.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Noback M.A., Tepstra P., Holsappel S., Venema G., Bron S.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RA Wendrich T.M., Marahiel M.A.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertolo M.G., Bessieres P., Boletín A., Borcherdt S.,
 RA Bourlès R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.X., Codani J.J., Comerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Gallazzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulle M.F., Itaya M., Jones L.,
 RA Jorja B., Karamata D., Katsuhara Y., Kjaer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Königstein G., Krogh S., Kunano M.,
 RA Kunita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manel C., Médigue C.,
 RA Medina N., Meliádo R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogilwara A., Odega B., Park S.R.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pulic C., Purnelle B., Rapoport G., Ray M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F.,
 RA Sekiguchi J., Sekowska A., Serot S.J., Serró P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Yamakoshi A., Tanaka T., Tepstra P., Tognoni A.,
 RA Tosiato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzsaecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis."
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X96983; CA65694.1; -
 DR EMBL: U58859; AAB01345.1; -
 DR EMBL: Z99108; CAB12739.1; -
 DR PIR: F69822; F69822.
 DR Subtilist; BG11588; yhcJ.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_lipoprot_9.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Membrane; Lipoprotein; Signal;

CC FT SIGNAL 1 19 POTENTIAL.

CC FT CHAIN 20 263 HYPOHETICAL LIPOPROTEIN YHC.

CC FT LIPID 20 20 N-palmitoyl cysteine (Potential).

CC FT LIPID 20 20 S-diacetylglycerol cysteine (Potential).

CC FT CONFLICT 79 79 A -> T (IN REF. 2).

CC FT CONFLICT 118 118 I -> L (IN REF. 2).

CC FT CONFLICT 170 170 H -> L (IN REF. 2).

CC FT CONFLICT 259 260 RL -> PA (IN REF. 2).

CC SQ SEQUENCE 263 AA; 29186 MW; 0513CTD91CDD2BFB3 CRC64;

CC Query Match 21.4%; Score 236; DB 1; Length 263;

CC Best Local Similarity 28.6%; Pred. No. 5e-13;

CC Matches 78; Conservative 70; Mismatches 99; Indels 26; Gaps 9;

Dy 9 ICA--LACGILAGCSNOSNEPALSKTAACITKGVWMAEGPQAAEYVAGYAKERKNTL 66

Dy 6 ICSTVLVLIVSFPTAS-----PSA---EHESIKIGI-AESDGAIWNYYIAQAEBA-GLD 54

Dy 67 VELVEFYNDYAMPNSAVSKEGLDANAMQHKKPYLEKDSKGGNLNVITVGTFEPYPAGYST 126

Dy 55 IQLPFSDYSASDLALANKEDIANAFQIIISYGFSTERYKL-NLAPLGTTTYTPMGITSK 113

Dy 127 KITKLINELKGATTAVPNDPSNLAPALLILEKOGIKIKDKNTNLFSTLIDVENPKLVT 186

Dy 114 RYERIRDISRGAVVSVDPKPAFDGFRALTTLVCGAGLLTLLKNGFNFGSGVDIMDNPRHLK 173

Dy 187 KEVDTSVARADIDVDLVANNVNNGYGUGL-----TASENGVFVEDKOSPYNITVARAD 241

Dy 174 KAVRQQ--DAVSGADVFWMKPSEKKKGLPKRTLTLSGGLMSDEE---MNLIVRAE 226

Dy 242 NKDSKALODPVKAYQTDEVEAEAKKFQDVIX 274

Dy 227 DQDFREALQTIELVQADDTAFFIEKEYOGDIVR 259

RESULT 16

RBP2_PLAVB

ID_RBP2_PLAVB STANDARD; PRT; 2867 AA.

QC Q00799; Q9NMZ3;

DT 01-APR-1993 (Rel. 25, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Retículoocyte binding protein 2 precursor (PVRBP-2).

GN RBP-2 OR RBP2.

OS Plasmodium vivax (strain Belen).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=31273;

[1]

RP SEQUENCE FROM N.A., AND REVISIONS TO 2438-2439.

RX MEDLINE=20299192; PubMed=10838229;

RA Galinski M.R., Xu W., Barmwell J.W.,

RT "Plasmodium vivax reticuloocyte binding protein-2 (PVRBP-2) shares structural features with PvRBP-1 and the Plasmodium yoelii 235 kDa thoptry protein family.";

RL Mol. Biochem. Parasitol. 108:257-262(2000).

[2]

RN SEQUENCE OF 1189-2439 FROM N.A.

RP MEDLINE=92315318; PubMed=1617731;

RX Galinski M.R., Medina C.C., Ingravallo P., Barmwell J.W.,

RA "A reticuloocyte-binding protein complex of Plasmodium vivax merozoites.";

RL Cell 69:1213-1226(1992).

-1- FUNCTION: INVOLVED IN reticuloocyte adhesion. Specifically binds to human reticuloocyte cells.

-1- SUBCELLULAR LOCATION: Membrane-bound (Probable).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

CC or send an email to license@sb-sib.ch).
-----
DR DR EMBL: AF184623; AA76525.1; -.
KW HSP; P03069; IGM.
FT SIGNAL 1 21
FT CHAIN 22 2867
FT DOMAIN 22 2805
FT TRANSSEM 2806 2826
FT DOMAIN 2827 2867
FT DOMAIN 44 133
FT DOMAIN 560 758
FT DOMAIN 1112 1285
FT DOMAIN 2758 2785
FT REPEAT 2758 2761
FT REPEAT 2762 2765
FT REPEAT 2766 2769
FT REPEAT 2770 2773
FT REPEAT 2774 2777
FT REPEAT 2778 2781
FT REPEAT 2782 2785
SQ SEQUENCE 2867 AA; 331433 MW; 6E7D8CA71AFBFD3 CRC64;

Query Match 8.1%; Score 112.5; DB 1; Length 2867;
Best Local Similarity 22.7%; Pred. No. 18;
Matches 65; Conservative 50; Mismatches 112; Indels 59; Gaps 13;

QY 33 KTAAGT-----KGVWAGEPQAVAVAGVAKKXNLVLELVEFNDYAMPNAAVS 83
DB 1903 KKKAKITLGVVDTSAKYVGKIKITPELATTELG-D-ATKTAQELFKSSKNVULETENS 1961
QY 84 K-GELDANAMQHKKPYL-----EKDSOEKGLNLIYVNTF--YYPIAGYGTKI 128
DB 1962 KNTNELDVHKNIDQAYKVALEILAHSDPIDTKQKDSKLTLEMGNQIYLVKLINQYKTKI 2021
QY 129 KITNELKDGAGTIAPV---DPSNLARA-----LILEKQGLIKLKONTNLFSTLD 176
DB 2022 SSISKSKEEAVSVKIGNVSKSHSELSKTSKDSKYDNILALEKQ--TELQNLNRSFTGKT 2079
QY 177 IVENPKGLVKEVDTSVVARAIDVDVLAVNNVYAGVGL--TASRGVFEVDKDSPEVN 234
DB 2080 NTNSDSGLKIKTIDFESLKNAKLTLE-----GVNALTKASDNDHEHVQSKSEP-VN 2129
QY 235 IIVARADNKD-----SKAIQDFVKAQVQDVEAEAKKQFRDGYIK 274
DB 2130 PALSEIEKERTIDISLNTALDELTKKGTCEV--SRYKLTKDTVTK 2173

RESULT 17
GTF2 STRDO
ID GTF2_STRDO STANDARD; PRT; 1592 AA.
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=9113227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase)".
RL J. Bacteriol. 173:589-596(1991).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

```


[2]
 RN SEQUENCE OF 1-60 AND 466-543 FROM N.A.
 RP STRAIN=ATCC 33530 /G-37;
 RC MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bolt K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 sequencing";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 proper assembly of unfolded polypeptides generated under stress
 conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 7 subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39721; AAC71620.1;
 DR EMBL: U02252; AAD12515.1; ALT_INIT.
 DR EMBL: U02268; AAD12534.1;
 DR PIR: D64243; D64243.
 DR HSSP: P06139; IGR1.
 DR TIGR: MG392;
 DR HAMAP: MF_00600; -; 1.
 DR InterPro: IPR001844; Chaperonin Cpn60.
 DR InterPro: IPR002423; Cpn60/TCP-1.
 DR InterPro: IPR008950; GroEL-ATPase.
 DR Pfam: PF00118; Cpn60 TCP1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone; ATP-binding; Complete proteome.
 KW SEQUENCE 543 AA; 58354 MW; 80FA5C7037C4BA8 CRC64;
 SQ

Query Match 7.5%; Score 103.5; DB 1; Length 543;
 Best Local Similarity 20.8%; Pred. No. 9.3; Indels 103; Gaps 15;
 Matches 70; Conservative 55; Mismatches 109;

QY 3 FGR-----INGICALASGIAL-AGCSNQ-----SNEPAISK-----TAAQ 37
 DB 7 FGRDARTLLQGIINKANAVKTVGKQNVLEKFNAPLLTNGCVTLAKIELSDPE 66
 QY 38 TIKVGMADPEQAVNAVAGQ-----VAKKKNLTVELVENNDYAMNSAVSKELDAN 90
 DB 67 NIGAKVISAVAASVTNDADGTTTALILAQEMTNRIETII--NKGANPNV--IRRGIEDAS 123
 QY 91 AMOHKPYLEKDSQSEKLANLTVIGNTFVYPLAGYSTKIKTLNELDGATIAVNDPSN-- 148
 DB 124 LLIRK-LRK-----YKKNITNEIRIQVAAISSGSEIKGL 159
 QY 149 LARALILKEKGLIKKNDNTLPSITLD-----IYENPKK 183
 DB 160 IQANMALVKNQVYI--TTDAKTIINTLTETEGIEFGKTVASPYNVDSEKMEVLEQPKI 218
 QY 184 LV-----IKEDNSVVARAIDVDVLAVYNNNVAGOVLTASENGVFEDSDSYNNII 236
 DB 219 IVSSLKINIKIILPLLESGVENGNPLIIVADPFAEVVTTTAAV-----KLRGTINAV 272
 QY 237 VARAD--NKDSKAIQDFYKAVQTVDEVEAKKQFSD 270
 DB 273 AVKCNVGEROKALEDLAISSGTLAVNTEINSQFSD 309

RESULT 20
 GTF1_STRDO STANDARD; PRT; 1597 AA.

AC P11001;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFI.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MFE28;
 RX MEDLINE=87308014; PubMed=3040686;
 RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
 RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
 sobrinus MFE28";
 RL J. Bacteriol. 169:4271-4278 (1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 CC -1- SIMILARITY: Belongs to family 70 of glucosyl hydrolases.
 CC -1- SIMILARITY: Contains 19 cell wall binding repeats.

DR EMBL: M17391; AAC63063.1;
 DR InterPro: IPR002479; CW-binding.
 DR InterPro: IPR003118; Glyco_hydro_70.
 DR Pfam: PF01473; CW-binding_1; 16.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
 KW SIGNAL
 FT CHAIN 1 38
 FT DOMAIN 39 1597
 FT DOMAIN 39 1050
 FT DOMAIN 1099 1597
 FT DOMAIN 1099 1597
 FT REPEAT 1099 1132
 FT REPEAT 1163 1213
 FT REPEAT 1227 1277
 FT REPEAT 1292 1342
 FT REPEAT 1352 1399
 FT REPEAT 1406 1455
 FT REPEAT 1465 1512
 FT REPEAT 1519 1568
 FT REPEAT 1582 1597
 SQ SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64;

Query Match 7.4%; Score 103; DB 1; Length 1597;
 Best Local Similarity 26.6%; Pred. No. 38;
 Matches 63; Conservative 28; Mismatches 90; Indels 56; Gaps 12;

QY 76 AMPSAVSKGELDANMOHPEYLEK-----DSQKGLNNTVIGNTFVYPLAGYSTIKTL 131
 DB 672 AMQVYQINGEILTVAYYGALKQSDKGDATRTTSVGVVMGQPNFSDG---KVAL 728
 QY 132 N-----ELKDG-ATIAVNDPSNIALALILKEKGLIKLKD-NTNLFSTT 174

Db 729 NMGAHANEDEYRALMNSTVDGATVATDADAS-----KAGLVKTEDENGYLYFLIN 778

Qy 175 LDI--VENPKQVKEVDTSVAPARAIDVDVAVNNVAGOVGLT-----ASENGVFE- 226

Db 779 DDLKGVANPQ--VSGFLQWVVPVGAADDODIRVAVASDTASTGKSIHQCAPAADSRYMFEQ 836

Qy 227 -----DKDSFYVITIA-RADNKDSKAIDPFVAYQTDVEAEAKKQFDGYIK 274

Db 837 FSNFSGFATKEEYITVAVIANNVDFKFSWGITDFENAPQ---YVSTGQFIDSVIQ 890

RESULT 21

AC APEA_CLOAB STANDARD; PRT; 465 AA.

AC Q97K30;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable M18-family aminopeptidase 1 (EC 3.4.11.-).

GN APEA OR CAC1091.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=1146286;

RA Noelting U., Berton G., Omeichenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R., Lee H.M., Dubois J., Qiu D., Hilti J., Wolf Y.T.,

RA Tatunov R.L., Sabatche F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.,

RT "Genome sequence and comparative analysis of the solvent-producing

RT bacterium Clostridium acetobutylicum."

RL J. Bacteriol. 183:4823-4838(2001).

CC -1- SIMILARITY: Belongs to peptidase family M18.

CC -1- COFACTOR: Zinc (By similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE007625; AAK79065.1; -

DR PIR: P97034; P97034.

DR HANAP; MF_00466;-1.

DR InterPro: IPR001948; Peptidase_M18.

DR Pfam: PF02127; Peptidase_M18; 1.

DR PRINTS: PRO0932; AMINO1PTASE.

KM Hypothetical protein; Hydrolase; Aminopeptidase; Metalloprotease;

KM Zinc; Complete proteome.

FT METAL 105 105 ZINC (POTENTIAL).

FT METAL 180 180 ZINC (POTENTIAL).

FT METAL 441 441 ZINC (POTENTIAL).

SO SEQUENCE 465 AA; 52011 MW; CB4C6D0A54C0A439 CRC64;

Query Match 7.2%; Score 99.5; DB 1; Length 465;

Best Local Similarity 22.6%; Pred. No. 14;

Matches 52; Conservative 39; Mismatches 75; Indels 65; Gaps 11;

Qy 77 MPNSAVSKGBLDANAMQKRYLEKDSQEKLNVLVGNFVPLAGYSTK-----IK 129

Db 1 MPNDLTK-----EYKNAMDK-YDDKQKKEVFALGDFKNFISNCTRECEYELK 50

Qy 130 TLNE-----LKDGATTAVPNDPSNLARALILBKQG-----IKLKONTNLF 171

Db 51 TAEKSGVRNEDILAKETLKEGDKYANRKGKL--IMPLIKELLYTGFKLL-GAHD 107

Qy 172 STLDIVENPKVLVKEVDTSVAPARALDDVLA VNNNTAG-----QVGLTASENGVIV 225

Db 108 SPRLDLKQNP-----LYEDTDLAMLETHYVGIKKYQWTLPLAIGHVIV 152

Qy 226 EDKDSFYVITIAVADNKDSKAIDPFVAYQTDVEAEAKKQFDGYIK 275

Db 153 K-XDGTIVNVCVGEDNDPVGVSIDILVHLASOLEKKASK-----VLEG 196

RESULT 22

AC SLAI_BACAA STANDARD; PRT; 814 AA.

AC P49051;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE S-layer protein sap precursor (Surface layer protein) (Surface array

DE protein).

GN SAP OR BA0895.

OS Bacillus anthracis (strain Ames), and

OS Bacillus anthracis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=18094, 1392;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Ames;

RX MEDLINE=22608414; PubMed=12721629;

RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,

RA Nelson K.E., Tetteijn H., Fouts D.E., Eisen J.A., Gill S.R.,

RA Holtzapfel E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,

RA Kolonay J.F., Beaman M.J., Dodson R.J., Binkac L.M., Gwin M.,

RA Desoy R.T., Madun R., Daugherty S.C., Durkin A.S., Haft D.H.,

RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,

RA Berton J.U., Mammond Y., Jiang L., Hance I.R., Weidman J.F.,

RA Berry K.J., Platt R.D., Wolf A.M., Watkins K.U., Nierman W.C.,

RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,

RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,

RA Fraser C.M.,

RT "The genome sequence of Bacillus anthracis Ames and comparison to

RT closely related bacteria."

RL Nature 423:81-86(2003).

CC -----

DR EMBL: AE017027; AAP24883.1; -

DR PIR: I40048; I40048.

DR TIGR: BA0885;-1.

DR InterPro: IPR003343; Big_2.

DR InterPro: IPR008964; Invasin_intimin.

DR InterPro: IPR001119; SLH.

DR Pfam: PF02368; Big_2; 1.

DR Pfam: PF00395; SLH_3.

KM Signal; Glycoprotein; Repeat; Cell wall; S-layer; Complete proteome.

FT SIGNAL 1 29

FT CHAIN 30 814 S-LAYER PROTEIN SAP.
 FT DOMAIN 33 93 SLH 1.
 FT DOMAIN 94 154 SLH 2.
 FT DOMAIN 155 213 SLH 3.
 SQ SEQUENCE 814 AA; 86620 MW; C1638D26A1C6B101 CRC64;

Query Match 7.2%; Score 99; DB 1; Length 814;
 Best Local Similarity 24.0%; Pred. No. 31;
 Matches 67; Conservative 40; Mismatches 120; Indels 52; Gaps 12;

QY 30 AISKRAQIKYGVAGPEQA--VAEVAAGVAKEKNLTYE-----LVENDYMPKSAVK 84
 DB 232 AVEKTKEDIKVTNKANDKVLKVEY--LSDDKSAVELYSNLAQKQTVDVNKKVK 289
 QY 85 GILDANAMQKRELEKD-----SOEKLNNLVIVNTFVYPLA----- 122
 DB 290 TEVAAGSLEAKTIEVADQTVVADEPTALQFTVKDNGNEVVSPEGIEFTVPAEKINAKG 349
 QY 123 -----GVSTKIKTNELKDQATVAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLLD 176
 DB 350 EITLAKGSTITVKAAYK-KDGKVAESKEKVSAGAAVASISNMTVAEONKADP-TSKD 407
 QY 177 IYENPKYVKEVDPSVVARAIDVDLAVNNNVAQVGLTASENGVEVEDKSPYVNI 236
 DB 408 FQONK--VYEGDAVYQVELKDQPNVY--TGKVEYESLNTFVAVDKATSKVTVL 460
 QY 237 VA-----RADNDSKAIDFVKAQYQTEVEAEAKKQFD 270
 DB 461 SAGKAPVKVTVKDSKG-KELVS--KTVEIEAPAKMKE 496

RESULT 23
 ID HXA2_HAEMIN STANDARD; PRT; 928 AA.
 AC P45354;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Heme/hemopexin-binding protein precursor (Heme:hemopexin utilization protein A).
 GN HXDA.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 NCBI_TaxID=727;
 [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-35.
 RC STRAIN=DL42 / Serotype B;
 RX MEDLINE=9515556; PubMed=7815944;
 RA Cope L.D., Thomas S.E., Latimer J.L., Slaughter C.A.,
 RA Mueller-Eberhard U., Hansen E.J.;
 RT "The 100 kDa haem-hemopexin-binding protein of Haemophilus influenzae: structure and localization.";
 RL Mol. Microbiol. 13:863-873(1994).
 RN (2)
 RP SEQUENCE OF 1-30 FROM N.A.
 RC STRAIN=DL42 / Serotype B;
 RX MEDLINE=95270579; PubMed=7751272;
 RA Cope L.D., Yoger R., Mueller-Eberhard U., Hansen E.J.;
 RT "A gene cluster involved in the utilization of both free heme and heme-hemopexin by Haemophilus influenzae type b.";
 RL J. Bacteriol. 177:2644-2653(1995).
 CC -1- FUNCTION: BINDS HEME/HEMOPEXIN COMPLEXES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; U08348; AAA74138.1; -.
 DR PIR; S54699; S54699.
 DR InterPro; IPR008638; Haemaggl. act.
 DR Pfam; PF05860; Haemaggl. act.; I.
 KW Transport; Signal; Repeat.
 FT CHAIN 1 21
 FT SIGNAL 1 21
 FT CHAIN 22 928 HEME/HEMOPEXIN-BINDING PROTEIN.
 FT DOMAIN 101 679 6 X 6 AA APPROXIMATE REPEATS.
 FT REPEAT 101 106 1-1.
 FT REPEAT 205 210 1-2.
 FT REPEAT 279 284 1-3.
 FT REPEAT 410 415 1-3.
 FT REPEAT 635 640 1-4.
 FT REPEAT 674 679 1-5.
 FT DOMAIN 149 172 4 X 6 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 149 154 2-1.
 FT REPEAT 155 160 2-2.
 FT REPEAT 161 166 2-3.
 FT REPEAT 167 172 2-4.
 SQ SEQUENCE 928 AA; 101228 MW; 67D4546A4B92390 CRC64;

Query Match 7.2%; Score 99; DB 1; Length 928;
 Best Local Similarity 23.7%; Pred. No. 36;
 Matches 62; Conservative 40; Mismatches 86; Indels 74; Gaps 13;

QY 5 KINGICALASGIALAGCSNOSNEPAISKTA-AQIKVGVAGPEQAQVAEVAQVAKEXY 63
 DB 119 EINVAAGLATTCKLERISENSNSYQFTRTKDQVLEKGLVADQGVKE--GVINEG- 175
 QY 64 NLTVELVEFNDYMPN--SAVSKGELDANAMQKRELEKQSKGLNNLVIVNTFVYPL 121
 DB 176 NITRQ-----DFVLANDEVINKNTN-----VEKSTNGK-----VYLS 211
 QY 122 AGYSTKIKTNELKDQATVAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLLDIVENP 181
 DB 212 SGVN-----FTFTLPDSGISVA-----LEDN-----TVQGIKVE 241
 QY 182 KLVYKREVDPSVAR--AIDVDVLAVNNNVAQVGLTASENGVEVEDKSPYV-- 234
 DB 242 GSIRAGEITLSAGKQALD---SLVNNGVLEATKVSNGKVVLSADNVELNNSNT 297
 QY 235 ---IVRADNDSKAIDFVK 253
 DB 298 KGEIVTFGADVTSNKEKDNK 319

RESULT 24
 ID TRME_CLOTE STANDARD; PRT; 459 AA.
 AC Q89982;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE tRNA modification GTPase tme.
 GN TRME OR THDF OR CTC00098.
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 NCBI_TaxID=1513;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Massachusetts / 588;
 RX MEDLINE=22457253; PubMed=12552129;
 RA Brueggemann H., Baesumer S., Fricke W.F., Wierer A., Liesegang H.,
 RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
 RA Gottschalk G.;
 RT "The genome sequence of Clostridium tetani, the causative agent of tetanus disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
 CC -1- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate. Involved in the biosynthesis of the hypermodified nucleoside 5-methylaminomethyl-2-thiouridine, which is found in the wobble position of some tRNAs (By similarity).


```

CC -----
DR EMBL; X59771; CAA42442.1; -
DR PIR; S1530; PCSOAG.
DR InterPro; IPR004829; Surface antigen.
DR InterPro; IPR005877; GspF_YsIRK.
DR InterPro; IPR001889; Gram_pos_anchor.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007756; RICH.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; YsIRK signal; 1.
DR ProDom; PD153432; Surface_antigen; 1.
DR SMART; SM00409; Ig; 1.
DR TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR TIGRfams; TIGR01168; YsIRK signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR PROSITE; PS00835; IG_LIKE; FALSE NEG.
DR Cell wall; Peptidoglycan-anchor; Receptor; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 37
FT CHAIN 38 1135 IGA FC RECEPTOR.
FT PROPEP 1136 1164 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 434 534 IG-LIKE.
FT DOMAIN 199 438 IGA-BINDING (POTENTIAL).
FT DOMAIN 439 826 IGA-BINDING (POTENTIAL).
FT DOMAIN 827 944 PRO-RTCH.
FT SITE 1132 1136 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1135 1135 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1164 AA; 131051 MW; 65DE94AF720A5474 CRC64;

Query Match
Best Match Similarity 7.0%; Score 97.5; DB 1; Length 1164;
Matches 67; Conservative 38; Mismatches 110; Indels 83; Gaps 14;

QY 8 GICALASGIALAGCSNQSNEPAISKTAQITKVGWAGPEQAVAEVAGQVAKKXNTL- 66
DB 20 GAASAVASALFPGSAVAHASE--LVKDDSVKTEVAAPKPSVAQDDQNNSSSLETTX 77
QY 67 -----VEIVFENDYAMPNSAVSGELDANAAMOHKPYLEDSQSEKGLNVLVIGN 115
DB 78 MELPTTDIKAVPEVE--KTAGEISATIDTGKSEKOLQWKNKLKD-----VDN 124
QY 116 TTVYPLAGYSTKIKTLNELKDGATIAVPDPSNLARALILEKQ-----GIKLDNTNL 170
DB 125 T-----ILSHQK-----NEFK-----TKIDETNDSALLEENQNETNRLHIKOH--- 167
QY 171 FETLIDIVENPKKL-----VIKEVDTSVAARAIDVDVLAVN---NNVAGQVGLTASENGV 223
DB 168 -----BEVEKDKKAAQOKTKLKQSDTK-----VDLSNIDKELNHQKSQVEKMAEQKGI 214
QY 224 FVEDKDSPYVNIIVAR-----ADNKDSKAIQ-----DFVKAAYQDVEAEAKKQ 267
DB 215 TVEDKDSMKIEDIIRKQAOQADKKEDAEVKYREBLGKLFSTYKAGLDQEIQEHVKE 272

RESULT 28
Y414 MYCGE STANDARD; PRT; 1036 AA.
AC P47653; P47654; Q49457;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG414.
GN MG414.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]_TaxID=2097;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,

```

```

RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fritchman J.L.,
RA Nguyen D.T., Uterback T.R., Saudak D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,
RA "The minimal genome complement of Mycoplasma genitalium.",
RT Science 270:1397-403 (1993).
RN [2]
RP SEQUENCE OF 52-146 AND 733-833 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RA "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.",
RT J. Bacteriol. 175:7918-7930 (1993).
RL [2]
RL -1- SIMILARITY: BELONGS TO THE MG414 / MG415 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL databank.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U39723; AAC71641.1; -
DR EMBL; U01695; AAB01008.1; -
DR EMBL; U01804; AAD12330.1; -
DR PIR; H64245; H64245.
DR TIGR; MG414; -
DR Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 4 24
FT TRANSWEM 1004 1024 LEPT-->SRAS (IN REF. 2)
FT CONFLICT 733 736
SQ SEQUENCE 1036 AA; 123179 MW; 12A21F00F686A141 CRC64;

Query Match
Best Match Similarity 7.0%; Score 97; DB 1; Length 1036;
Matches 62; Conservative 43; Mismatches 91; Indels 116; Gaps 13;

QY 71 EENDY-----AMPSAASKG---ELDANA-----MHKPYLEDSQSEKGL- 107
DB 687 KENDYERKQWAFITKNSFKSKDKNDIDLINSFLFVNDPKKYPKALPTNEKKGIFVYL 746
QY 108 -----NNTL-----VIVGNFVYPLAGYSTKIKTL-----NEV----- 134
DB 747 AEINNSNELFKYGSSESDSEIINDKQVFLSQNKPNPLRYLPKPHNTKLFDVNDGFKL 806
QY 135 -KDGATIAVPN-----DPSNLARALILEKQGL----- 161
DB 807 EKDKVPLVDNKLKIALDLKSSFKFLEDYQDLDFHFPFLINDQQLVLDLNTLSEKRLQ 866
QY 162 -----TKLDNTNLFTLIDIVENPKKLVIKEVDTSVAARAIDVDVLAVV 206
DB 867 TTGNVAFNLKKEPNIHVLVENVKQFVLVDVVRSKGLFIKGVNNDNVFSI-SYDLKTT 925
QY 207 NNNVAGQVGLTASENGFVEDKDSPYVNIIVA-----RADNKDSKAIQDF-----VKAYQ 256
DB 926 NNQTLTIVANGFNDIWFIDTISENQTLFKALSFYKQNNLQFRRVDFNLKSODKSYE 985
QY 257 TDEVEA-EAKKQ 267
DB 986 VDKLEKNEIKQ 997

RESULT 29
PUR5 OCEIH STANDARD; PRT; 339 AA.
AC Q8ES94;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1) (AIRS)

```


DE (Phosphoribosyl-aminimidazole synthetase) (AIR synthase).
GN PUM OR 080747.
OS Oceanobacillus theysensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theysensis isolated from the Ithya Ridge and its unexpected adaptive capabilities to extreme environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
CC -1- CATALYTIC ACTIVITY: ATP + 2-(formamido)-N(1)-(5-phospho-D-ribose)1-imidazole = ADP + phosphate + 5-amino-1-(5-phospho-D-ribose)1-imidazole.
CC -1- PATHWAY: De novo purine biosynthesis; fifth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the AIR synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP004595; BAC12703.1; -
DR HAWAP; MF_00741; -; 1.
DR InterPro; IPR000728; AIR_synth.
DR InterPro; IPR004733; PurM_c1igase.
DR Pfam; PF00586; AIRS; 1.
DR Pfam; PF02769; AIRS_C; 1.
DR TIGRfams; TIGR00878; purM; 1.
DR Purine biosynthesis; Ligase; Complete proteome.
KM Purine biosynthesis; Ligase; 99DDA067B2445F CXC64;
SQ SEQUENCE 339 AA; 36326 MW; 99DDA067B2445F CXC64;
Query Match 7.0%; Score 96.5; DB 1; Length 339;
Best Local Similarity 20.6%; Pred. No. 15;
Matches 52; Conservative 41; Mismatches 78; Indels 81; Gaps 11;
QY 26 NEPAIASTACTIVGMAPEQA-----VAEVAQVAKKKNL---TYELVE--- 71
DB 111 NDPSRI-----EAVISGIAEGCEGAGALIGGETEMGMYDPPEYDLAGFVGVIEKSA 165
QY 72 -----FNDYAMPNSAVSGEL-----DANAMQKPY 97
DB 166 MITGDKISGVVITLSSGSHNSGYSILVKLIADVANDNQYVPLGSCVQKAVAPRTIY 225
QY 98 -----LEKDSQEKGLNNLVIVGNTFVPLAGYSTKIKITLNELDGATIANPDSNLAR 151
DB 226 AKSIQALKKXVNLKGIISH--ITG-----GGFDENIPRM--LPDGLVLLTETMSWLDIPE 274
QY 152 ALILEKGLILKIKNTNLFSTLTDIVENPKLVIKEDVTSVAARAIDVVLAVNNNYA 211
DB 275 VHFHLEBKNDINRENYGVFNWIGMA-----VVAEDVSLALQLLEKVD-----EQA 323
QY 212 GQVGLTASENGV 223
DB 324 VVIGKVTREGEV 335
RESULT 30
GLTI YEAST STANDARD; PRT; 2144 AA.
AC Q12680; Q12290;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Glutamate synthase [NADPH] precursor (EC 1.4.1.13) (NADPH-GOGAT).
GN GLTI OR YD1171C.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CN36;
RX MEDLINE=97082505; PubMed=8923741;
RA Flichef P., Varetgani M.P., Valenzuela L., Gonzalez A., Ballario P.;
RT "Sequence of the GLTI gene from Saccharomyces cerevisiae reveals the domain structure of yeast glutamate synthase."
RL Yeast 12:1359-1366(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=8288c;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 2 L-glutamate + NADP(+) = L-glutamine + 2-oxoglutarate + NADPH.
CC -1- COFACTOR: Binds a 3Fe-4S cluster; FAD and FMN.
CC -1- SUBUNIT: Homotrimer.
CC -1- SIMILARITY: TO OTHER GLUTAMATE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X88221; CA61505.1; -
DR EMBL; Z67750; CA61574.1; -
DR EMBL; Z74219; CA68745.1; -
DR GenOnline; 140414; -
DR SGD; S0002330; GLTI.
DR GO; GO:0005623; Cell; IDA.
DR GO; GO:0006537; P; Glutamate biosynthesis; IEP.
DR InterPro; IPR000785; Adnrdx_reductase.
DR InterPro; IPR002489; DUF14.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR006982; Glu_synth_cent.
DR InterPro; IPR006981; Glu_synth_NTN.
DR InterPro; IPR002932; Glu_synthase.
DR InterPro; IPR006005; Glu_synth_subl.
DR InterPro; IPR001003; Pyridine_redox_2.
DR Pfam; PF04898; Glu_synth_central; 1.
DR Pfam; PF04897; Glu_synth_NTN; 1.
DR Pfam; PF01645; Glu_synthase; 1.
DR Pfam; PF01493; GXGKG; 1.
DR Pfam; PF00070; Pyr_redox; 1.
DR PRINTS; PR00419; ADXRDPTASE.
DR PRINTS; PR00368; FADNR.
DR PRINTS; PR00469; PNDKPTASII.
DR TIGRfams; TIGR01317; GOGAT_sm_gam; 1.
KW Oligoeductase; Iron-sulfur; 3Fe-4S; Flavoprotein; FAD; FMN; NADP; Glutamate biosynthesis; Zymogen.
FT PROPEP 1 53
FT CHAIN 54 2144
FT NF_BIND 1131 1183
FT METAL 1184 1184
FT METAL 1190 1190
FT METAL 1195 1195
FT CONFLICT 30 30
FT CONFLICT 166 172
FT CONFLICT 449 451
FT CONFLICT 1752 1752
FT L -> D (IN REF. 2).
FT LPS -> ELV (IN REF. 2).
FT L -> V (IN REF. 2).
SQ SEQUENCE 2144 AA; 238200 MW; 5AA6A948EF995345 CXC64;
Query Match 7.0%; Score 96.5; DB 1; Length 2144;
Best Local Similarity 27.8%; Pred. No. 1,5e+02;
Matches 67; Conservative 24; Mismatches 89; Indels 61; Gaps 14;

QY 73 NDYAMNSAVSGELDANMOMKPYLEKDSQKGLNNLIVGNTFYPLAGYSTKI-----128
 Db 1377 NDY-----VGKQ-LSGGIIVIKP--FKDSKFKSDBN-VIVANTCFYATSGTAFISGSA 1426
 QY 129 -KTLELNEKQATIAVPNDPSNLA-----PALITLEKQGLIKLMDNTN-----LFSPTL 175
 Db 1427 GEFGRVGRNSGATIVVERIKGNNAFEYMTGGRALIVLSQMSLNAFSGATGIALYCLTSDYD 1486
 QY 176 DIVENPKULVKEVDTSVARAALIDVDLAIVNN-----NYAGQVGLTASENG-----222
 Db 1487 DFEVGIKIN-----DIVELESCLDPVEIAFVKNLLOEHNNY-OSDLBARILGNFNHYLK 1539
 QY 223 VFVEDKSDPYNNIV-----RADNKSKAIDDFKAVQ-----DDEVA-----BAKQF 268
 Db 1540 DFKVPTVTKYKVLKKEKKAAPAKAKAKATSEYTKFRSQEVDDEVNTILLANQAKQES 1599
 QY 269 K 269
 Db 1600 K 1600

RESULT 31
 V-ATPase HUMAN STANDARD; PRT; 382 AA.
 ID V-ATPase HUMAN STANDARD; PRT; 382 AA.
 AC P21283;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vacuolar ATP synthase subunit C (EC 3.6.3.14) (V-ATPase C subunit)
 GN ATP6V1C1 OR ATP6C OR V-ATPase C subunit.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Osteoclastoma;
 RA MEDLINE=94071935; PubMed=8250920;
 RA van Hille B., Vanek M., Richener H., Green J.R., Bilbe G.;
 RT "Cloning and tissue distribution of subunits C, D, and E of the human
 RL vacuolar H(+)-ATPase.";
 RD Biochem. Biophys. Res. Commun. 197;15-21(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA MEDLINE=21574584; PubMed=11707601;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
 RA Mrozek K., Sill H., Knutilla S., Kolitz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BALUC, the human member of a novel mammalian neuroectoderm gene
 RL lineage, is implicated in hematopoiesis and acute leukemia.";
 RD Proc. Natl. Acad. Sci. U.S.A. 98;13901-13906(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshitsugu S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.C.,
 RA Boesek S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whittinger R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakesley R.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Rodriguez A.C., Rodriguez J., Schmutz J., Skalska U., Smalins D.E.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99;16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 36-382 FROM N.A.
 RC TISSUE=Brain
 RX MEDLINE=91056085; PubMed=2147024;
 RA Nelson H., Mandiyan S., Nouni T., Moriyama Y., Miedel M.C.,
 RA Nelson N.;
 RT "Molecular cloning of cDNA encoding the C subunit of H(+)-ATPase from
 RT bovine chromaffin granules.";
 RL J. Biol. Chem. 265;20390-20393(1990).
 CC -1- FUNCTION: Subunit of the peripheral V1 complex of vacuolar ATPase.
 CC Subunit C is necessary for the assembly of the catalytic sector of
 CC the enzyme and is likely to have a specific function in its
 CC catalytic activity. V-ATPase is responsible for acidifying a
 CC variety of intracellular compartments in eukaryotic cells.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a
 CC peripheral catalytic V1 complex (components A to H) attached to
 CC an integral membrane V0 proton pore complex (components: a, c, c',
 CC c'', and d).
 CC -1- SIMILARITY: Belongs to the V-ATPase C subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; X69151; CAA48903.1; -;
 DR EMBL; AF363578; AAL50383.1; -;
 DR EMBL; BC010960; AAL10960.1; -;
 DR EMBL; J05682; AAA36803.1; -;
 DR PIR; JN0907; JN0907.
 DR Genew; HGNC:856; ATP6V1C1.
 DR MM; 603097; -;
 DR GO; GO:0016669; C:proton-transporting two-sector ATPase complex; TAS.
 DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. .; TAS.
 DR GO; GO:0015992; P:proton transport; TAS.
 DR InterPro; IPR004907; V-ATPase_C.
 DR Pfam; PF03223; V-ATPase_C; 1.
 DR KEGG; K01101; ATP synthase; ATP synthetase; Hydrogen ion transport.
 KM Hydrolyase; ATP synthetase; Hydrogen ion transport.
 SQ SEQUENCE 382 AA; 43941 MW; 5626E2AB2BD66BA7 CRC64;

Query Match 6.9%; Score 96; DB 1; Length 382;
 Best Local Similarity 23.0%; Pred. No. 19;
 Matches 70; Conservative 47; Mismatches 99; Indels 88; Gaps 16;

QY 18 LAGCSNSQNEPAAIKSTAQTIKVG---VMAGPEQAVAEVAGVAKKYNLVELEVFEND 74
 Db 22 LHAATSKNNNAVATSKFNIPDLKVGTLVAVG-----LSDELAKLD- 62
 QY 75 YAMPNSAVSKGELDANMOMKPYLEKDSQKGLNNLIVG-----NTFYPLAGYSTK 127
 Db 63 -AFVEGVVKK-----VAQYMDVIVEDSKQKQENMLANGVDLVYIRFGQDMKAKYPIK 115
 QY 128 -IKTLELNEKQATIAVPNDPSNLA-----PALITLEKQGLIKLMDNTN-----LFSPTL 175
 Db 116 QSKNLSITLIANGVTQIDDLKRSASAYNTL--KQNLONLEKQVAGSLTSLAE-----168
 QY 185 VIKEDVTSVARAALIDVDLAIV-----NNYAGQVGLTAS-----ENGVFEDKSDPYNN 234
 Db 169 IYKQDVFILDSYLL-VTLTLLVVPPLKLNNDWIKQYETLAEMVVPSSVVLSDQDSYLCN 226
 QY 235 IIVARAADKSDKALIDF-----VKAYOTDEVEEA-----KKQFKQGV 272
 Db 227 VTLFR-----KAVDFHKAENKFTIVRDQYNEEKAKDEENRSLTDKQKQF--GP 278
 QY 273 IKGM 276

Db 279 LVFW 282

RESULT 32

ARO_A_CAMJE STANDARD; PRT; 428 AA.

AC P52312; Q9PP36; 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPs).

GN AROA OR Cj0895C.

OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteres; Campylobacteraceae; Campylobacter.

OX NCBI_TaxId=197;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=81116;

RX MEDLINE=97128776; PubMed=8973316;

RA Woesten M.M.S.M., Dublink V.H.J., van der Zeijst B.A.M.;

RT "The aroA gene of Campylobacter jejuni.";

RL Gene 181:109-112(1996).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=NCIC 11168;

RX MEDLINE=20150912; PubMed=10688204;

RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C., Basham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S., Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W., Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M., Whitehead S., Barrett B.G.;

RA "The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.";

RT Nature 403:665-668(2000)

RL Nature 403:665-668(2000)

CC -1- CARBONYL ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate = phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.

CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway; sixth step.

CC -1- SUBUNIT: Monomer (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic. (Probable).

CC -1- SIMILARITY: Belongs to the EPSP synthase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL, X89371; CAA61554.1; -.

CC EMBL, AL139076; CAB73153.1; -.

CC PIR, G81362; G81362.

CC PIR, JCS338; JCS338.

CC HAMAP, MF_00210; -; 1.

CC InterPro, IPR006264; AROA.

CC InterPro, IPR01986; EPSP_synth.

CC Pfam, PF00275; EPSP_synthase; 1.

CC Prodom, PD001867; EPSP_synthase; 1.

CC TIGRfam, TIGR01356; aroA; 1.

CC PROSITE, PS00104; EPSP SYNTHASE 1; 1.

CC PROSITE, PS00885; EPSP SYNTHASE 2; 1.

CC Aromatic amino acid biosynthesis; transferase; Complete proteome.

CC CONFLICT 152 152 K -> N (IN REF. 1).

CC CONFLICT 158 158 F -> Y (IN REF. 1).

CC CONFLICT 179 179 D -> Y (IN REF. 1).

CC CONFLICT 183 183 T -> A (IN REF. 1).

CC CONFLICT 217 217 N -> S (IN REF. 1).

CC CONFLICT 241 241 V -> A (IN REF. 1).

CC CONFLICT 428 428 V -> A (IN REF. 1).

Query Match 6.9%; Score 96; DB 1; Length 428;

Best Local Similarity 22.1%; Pred. No. 22; Matches 62; Conservative 54; Mismatches 103; Indels 62; Gaps 14;

Db 8 GICALAG-IALAGCSQSNENP-AAISKTAQITIKVGMAPCAVAVGAVAKKYNL 65

99 GFLAGISGFVLSGDKLNRPKRISKPLTQ---IG-----ARIGRNANLPL 146

Cy 66 TVE---LVERNDYAMPNSAVSKGELDANMQ-----HKPYEKSGQ---KGLNVL 112

Db 147 CIEQOKLKARNFSEISSAQVKTAMILAFADNVCTFSEISIRHNSNNLAKMKAPR 206

Cy 113 VGNFTVYPLAGYSTKIKTLNELKDGATIAVNDPSN---LARAILLKQGLIKLQNT 168

Db 207 VUSD-----GLSRIPLKPKLKAQNTIIPNDPSAFYVLAAILDPSQITLK----- 255

Cy 169 NLFSTLDIVENPKKLVKEVDTSVARAIDVDLAVNNV---AAQVGLTASE-NGVFL 225

Db 256 -----NILNLPRIEAYKILQMGAK---LEMTITQDFTETIGIRVSSKNGIEV 304

Cy 226 EDK-----DSPYNTIVARADNKKSKAIDPFVKAQOTDEV 260

Db 305 KDNITMLIDEPALALAFALAKGSSLINAKELRVESDRI 345

RESULT 33

SPL_RARPA STANDARD; PRT; 525 AA.

AC Q05308; 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine protease I precursor (EC 3.4.21.-) (Rpi).

OS Rarobacter faecitabidus.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococciaceae; Rarobacteraceae; Rarobacter.

OX NCBI_TaxId=13243;

RN (1)

RP SEQUENCE FROM N.A., AND SEQUENCE OF 212-222; 224-238 AND 240-244.

RC STRAIN=YLM-50;

RX MEDLINE=93094226; PubMed=1339445;

RA Shimoi H., Iimura Y., Obata T., Tadenuma M.;

RT "Molecular structure of Rarobacter faecitabidus protease I. A yeast-lytic serine protease having mannose-binding activity.";

RL J. Biol. Chem. 267:25189-25195(1992).

RN (2)

RP SEQUENCE OF 212-247.

RX MEDLINE=92138668; PubMed=1778983;

RA Shimoi H., Tadenuma M.;

RT "Characterization of Rarobacter faecitabidus protease I, a yeast-lytic serine protease having mannose-binding activity.";

RL J. Biochem. 110:608-613(1991).

CC -1- FUNCTION: MAJOR SERINE PROTEASE EXHIBITING LYTIC ACTIVITY TOWARD LIVING YEAST CELLS. SIMILAR TO ELASTASE IN ITS SUBSTRATE SPECIFICITY AND HAS A TRYPTIC-LIKE APPARENTLY FOR MANNOSE. MANNOPROTEIN MAY BE THE NATIVE SUBSTRATE FOR RPI.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: Contains 1 ricin B-type lectin domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL, D10753; BAA01585.1; -.

CC PIR, A45053; A45053.

CC HSP, P00778; IGSA.

DR InterPro: IPR004236; AL protease.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR001254; peptidase S1.
DR InterPro: IPR001316; peptidase S1.
DR InterPro: IPR000772; Ricin B lectin.
DR InterPro: IPR008997; Ricin B-like.
DR Pfam: PF02983; AL protease; 1.
DR Pfam: PF00652; Ricin B lectin; 3.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00865; ALTYCPTASE.
DR SMART: SM00458; RICIN; 1.
DR PROSITE: PS00134; TRYPSIN_RHS; 1.
DR PROSITE: PS00135; TRYPSIN_LCTR; 1.
DR PROSITE: PS00231; Ricin B LECTIN; 1.
KW Hydroxylase; Serine protease; Mannose-binding; signal; zymogen;
KW Lectin.
FT SIGNAL 1 32
FT PROPEP 33 211
FT CHAIN 212 535
FT DOMAIN 336 535
FT DOMAIN 401 525
FT DISULFID 223 239
FT DISULFID 310 320
FT DISULFID 346 376
FT DISULFID 412 431
FT DISULFID 453 472
FT ACT SITE 238 238
FT ACT SITE 270 270
FT ACT SITE 352 352
SQ SEQUENCE 525 AA; 55654 MW; DA3BCFD330EBB61 CRC64;
Query Match 6.9%; Score 96; DB 1; Length 525;
Best Local Similarity 24.2%; Pred. No. 28; Indels 88; Gaps 14;
Matches 75; Conservative 30; Mismatches 117;
DR 11 ALASGIALAGCSNOSNEPAISKTAQITKVG-VNAGPEQAVAEVAGQVAKENYLIVTEL 69
DB 8 ALFSLALVGAAGASVGLAASANSAPAAATVQASSSATSVAAATSKQDGEVLA 67
QY 70 VERNDYAMNSAVSKG-ELDANAMQHPYLEKDSQEK--GL-----NNLVYGNTEF----- 117
DB 68 V--RDKITKTQAKRIKLEEKARQLEPRLQKKLQKKEFAGLWISKNGKTIYGVTTKRA 125
QY 118 -VYPLAGYSTK--KTLNELKQATLAVPNDPSNLARALLILEKQGLIKLQNTNLFSTT 174
DB 126 KVVKKAGATPKIVKSNLTLTKRAIKISKAPS-----DIXVNSW 167
QY 175 LDIENPKLVIKEVDTSVAPARAIDVDLAVNNVYAGVGLTASENGVEEDKDSPYVN 234
DB 168 VDPATN--KVI--EASRKAARAA-----ATAAGLTAGTYEITVSD-----D 206
QY 235 IYARADNDSKAIQDFVAY-----QDVEAEAKK 266
DB 207 VIVPVDVYGGALSGCTLAFPYVGGFLTAGCAVEKGHLIKTENTGGQIGTVEA---S 263
QY 267 QFKDGVIKGM 276
DB 264 QFGDGIIDAM 273
RESULT 34
ID CH60 AZOVI STANDARD; PRT; 546 AA.
AC ORGBA4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR GROEL.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UMR;
RA "an active group II intron has invaded the genus Azotobacter and is
RT inserted within the essential groEL gene";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ database.
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AY057439; AAL25964.1; -.
DR HAMAP: MF_00600; -1.
DR InterPro: IPR001844; Chaperonin Cpn60.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR InterPro: IPR008950; GroEL-ATPase.
DR Pfam: PF00118; Cpn60_TCP1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding.
KW SEQUENCE 546 AA; 56875 MW; 8E1F4C93E42A619 CRC64;
Query Match 6.9%; Score 96; DB 1; Length 546;
Best Local Similarity 20.8%; Pred. No. 30;
Matches 74; Conservative 49; Mismatches 102; Indels 130; Gaps 16;
QY 8 GICALSG-----IALAGCSNOSNEPAISKTAQ--TIKVGVAPEQAV 51
DB 103 GLAVVAAGNPMPLDKGIDKATITAVELKSLAKPCSDSKALQVTTIANSDESIGNII 162
QY 52 AEVAGVAKENYLIVTE-----LVY-----ENDYAMPNSAVSKGELD----- 88
DB 163 AEMNRYKRGKV--ITYEBSGLENELSVYEGMDFDGYLSPYRINRPTDVAELDNPILL 221
QY 89 -----ANAMQHPYLE-----KDSQEGANNVIVGNTFYPLA-----GY 124
DB 222 LVDKISINIRELLPVEAVAKSGRPLLIVAEDVEGALATLVNNMRGIKVAAYTAPGF 281
QY 125 STKIKTLNELKQATLAVPNDPSNLARALLILEKQGLIKLQNTNLFSTLDIENPKKL 184
DB 282 NDRKAMLYQ-----IAL-----LTGATVISKVL-----SLESATLEHLGKPKAL 323
QY 185 VIKEDVTSV--AAPAIDVDLAVNNVYAGVGLTASENGVEEDKDSPYVNIIVARADN 242
DB 324 VLNKENTTIMGAGAQDIEAAVAQIRKQ-----IEFTSSPY----- 360
QY 243 KDSKAIQD-----FYKAVQDVEAEAKK-----QFKDGVIKG 275
DB 361 -DRKLDQERLAKAGVAVIKVGAATVEKKEKKAAYEALAHATRAAYEGVVP 414
RESULT 35
ID RRP5 YEAST STANDARD; PRT; 1729 AA.
AC Q05022;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE RNA biogenesis protein RRP5.
GN RRP5 OR FMI1 OR YMR229C OR YMG959.11C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288 / AB972;
RC MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagsis K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skellern J., Walsh S., Whitehead S., Barrall B.G.,
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII."
RL Nature 387:90-93(1997).
RM [2]
RP CHARACTERIZATION.
RX MEDLINE=97051828; PubMed=8896463;
RA Venema J., Tollervey D.,
RT "RRP5 is required for formation of both 18S and 5.8S rRNA in yeast."
RL EMBL J. 15:5701-5714(1996).
CC -1- FUNCTION: INVOLVED IN THE BIOGENESIS OF RNA. REQUIRED FOR THE
CC FORMATION OF 18S AND 5.8S RNA.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -1- SIMILARITY: Contains 4 HAT repeats.
CC -1- SIMILARITY: Contains 11 SI motif domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 249939; CAA90200.1; -
DR PIR: S57596; S57596.
DR HSSP: P05055; ISRO.
DR Germline: 142904; -
DR SGD: S0004742; RRP5.
DR GO: GO:0005731; C:nucleolus organizer complex; IPI.
DR GO: GO:0030515; P:snRNA binding; IPI.
DR GO: GO:0030490; P:processing of 20S pre-rRNA; IPI.
DR InterPro: IPR003107; HAT.
DR InterPro: IPR008994; Nucleic_acid_OB.
DR InterPro: IPR003029; SI.
DR InterPro: IPR008941; TPR-like.
DR Pfam: PF00575; SI.7.
DR SMART: SM00386; HAT; 6.
DR SMART: SM00316; SI.12.
DR PROSITE: PS50126; SI.11.
KM Nuclear protein; rRNA processing; Repeat.
FT DOMAIN 115 200 SI MOTIF 1.
FT DOMAIN 338 410 SI MOTIF 2.
FT DOMAIN 510 580 SI MOTIF 3.
FT DOMAIN 607 676 SI MOTIF 4.
FT DOMAIN 690 769 SI MOTIF 5.
FT DOMAIN 794 863 SI MOTIF 6.
FT DOMAIN 885 971 SI MOTIF 7.
FT DOMAIN 1003 1083 SI MOTIF 8.
FT DOMAIN 1088 1159 SI MOTIF 9.
FT DOMAIN 1177 1245 SI MOTIF 10.
FT DOMAIN 1265 1336 SI MOTIF 11.
FT REPEAT 1455 1487 HAT 1.
FT REPEAT 1561 1594 HAT 2.
FT REPEAT 1632 1664 HAT 3.
FT REPEAT 1666 1701 HAT 4.
SQ SEQUENCE 1729 AA; 193133 MW; 39BF46E5587B3B0A CRC64;

Query Match 6.9%; Score 96; DB 1; Length 1729;
Best Local Similarity 21.0%; Pred No 1 2e-02;

Matches 57; Conservative 46; Mismatches 82; Indels 86; Gaps 14;
QY 77 MPNSAVSK-----GE-----LDANAMQH-----KPELRDSEKQKIANLVI 112
DB 634 LPNSEISEVFVKREHRLRGQIVYVLLVDADARRRIINATCKSNQNAQKQNTINIV 693
QY 113 VGNFTVFPPLAGYSTKIKITLNEBKDGATIAVFN-----DPSNLARAL 153
DB 694 PGRITII-----TVVIEKTKDVSIVEIPDVGLRGVIVYGHLSDRIRIQNRAQLMKLR 745
QY 154 ILLKQGLIKLKD-NTNLFTTL--DIVENPKLVI-----KEYDTSVAAAIID 199
DB 746 IGTLTGLVIDKDRTRTFVFNMSLKSILKAKKETLPLTDDVDKLNKDPMAIYKISIS 805
QY 200 DVDLAVNNNNAAG-VGLTASENGVEVEDKD--SPYVN---IIVRADNKSQALQDF 251
DB 806 DKGLFVAFN--GKFIGLVLPSYAVDSRDIDISKAFYINQSVYVYLLRTDDKQK----F 858
QY 252 VKAYQTEVEAEAKQK-----DVYIKGM 276
DB 859 LLSLKAPKVEKKKVESNIEDPVSSIKSM 889
RESULT 36
ID AROA AOUAE STANDARD; PRT; 431 AA.
AC 067494;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPPS synthase) (EPPS).
GN AROA OR A0_1536.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aulay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."
RL Nature 397:353-358(1998).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC sixth step.
CC -1- SUBUNIT: Monomer (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the EPPS synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000744; AAC07443.1; -
DR PIR: D70433; D70433.
DR HAMAP: MF_00210; -; 1.
DR InterPro: IPR006264; AROA.
DR InterPro: IPR001986; EPPS synth.
DR Pfam: PF00275; EPPS synthase; 1.
DR ProDom: PD001867; EPPS synthase; 1.
DR TIGRFAMs: TIGR01356; aroA.1.
DR PROSITE: PS00104; EPPS SYNTHASE 1; 1.
DR PROSITE: PS00885; EPPS SYNTHASE 2; 1.
KM Aromatic amino acid biosynthesis; Transferrase; Complete proteome

Query Match	Best Local Similarity	6.9%; Score 95.5; DB 1; Length 684;
Matches 60;	Conservative 19.4%; Pred. No. 42;	
	Mismatches 53;	Indels 118; Gaps 79; Gaps 14;
QY 4	KGINGICAL	ASGIALAGCSNSQSEPPAISTKAQTI--KGVWA-----GPEQA 50
DB 329	GKLGWVSLNQCKIDL	SALTSESTSEBNDTALIIQAINAKGLAFNNAEGKLVINSKTG 388
QY 51	VAAEYAGVAKKXNLT	VELVEFNQYAPMNSAVSKGELDANAMQKPYLEKDSQEKGINLL 110
DB 389	MLTIKGEHLAKSL	-----KDLGNAGVQSEVSAQNTLFPMSKULQKASD----- 434
QY 111	VIVGNTFPYPIAGY	STIKITNELKQGTAV-----PNDSNLARALLILEKQGLTK 165
DB 435	---SAFTYNGVS	ITRPTNEVNDVIGVNITLLEOTTEPNKR-----AIISV-----R 478
QY 166	DNTMLFSTLIDV	NPKKLVK-----EVDTSVAA--RAIDVDVL--AVANNVYAGV 214
DB 479	DNQAIIDSLT	FPVANNELLIKLDEDRYDADTKIAGFNQVGIIRARSSLNVFSYSV 538
QY 215	-----GLTASENG	VEVEDKDSPYVNIIVAPADNKSQALQDVFAYQTDVEYA 262
DB 539	HTDNGVESLMKYG	SLDKDQWMSIDEAK-----LSLSALNSNPATODFF--YGSDSXDM 590
QY 263	EAKQKQFKDGV	272
DB 591	GGRRIHQEGI	600
<p>RESULT 38 SCA4_RICR ID_SCA4_RICR STANDARD; PRT; 1013 AA.</p> <p>AC 99AD81; DT 16-OCT-2001 (Rel. 40, Created) DT 16-OCT-2001 (Rel. 40, Last sequence update) DT 16-OCT-2001 (Rel. 40, Last annotation update) DE Antigenic heat-stable 120 kDa protein (PS120) (120 kDa antigen) DE (Protein PS 120) (Fragment). GN SCA4 OR D. OC Rickettsia rhipicephali; OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;</p>		

CC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=33392;
RN [1]
RP SEQUENCE FROM N.A.
RA Sekoyova Z., Roux V., Raoult D.;
RT "Phylogenetic analysis of Rickettsia spp. by comparing sequence of the
RL 'gene D' coding for an intracytoplasmic protein."
Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF15053; AAK30684.1; -
CC Antigen.
KM NON TER
FT NON TER
SQ SEQUENCE 1013 AA; 110550 MW; 856E98912315D102 CRC64;

Query Match 6.9%; Score 95.5; DB 1; Length 1013;
Best Local Similarity 20.3%; Pred. No. 69;
Matches 59; Conservative 40; Mismatches 124; Indels 67; Gaps 9;

QY 31 ISKTAQITKVGWVAGPEQAAVAVGQVAKKYNL-----TVELV 70
DB 641 LKTNIGQITSNVLDGP--ATREVGKSLIQTETVAVSSLEADQKAIYKVGETTAT 698
QY 71 EFNDAVMPNSAVSKGELDANAMQHPYV-EKDSCEKGLNIVVNTFVPLAGYSTIK 129
DB 699 SDTSLSLRKLIMASAEKGIYESTKTNLPDRMLMKGLVDGIYBEKSGPELTAVSSGD 758
QY 130 TLNEIKDQATIVPNDPSNLARALILKEKGLI-----KLKDN 170
DB 759 NSN-INDESEKALKKAKDAASEATLIDITONTLEGLKQONTIEHKPRDDIYKQAEVNA 817
QY 171 PSTTLDIYENPKKVIYKE-----VDTSVAARIDVDVAVNNNVAQVGLTASGVF- 224
DB 818 VNPVLEALEKPAVPSAERIVQETSSLNINISKALEKYNFFA-----MLSSNGPK 872
QY 225 -VEDKSPVYNIIVARADNKSKAIDPVKAYQDVEVAEAKKQPKDGYI 273
DB 873 TLEKKKEESI-----KKYDEIVKAFGTGKSTEEQGSFKAMLI 910

RESULT 39
HAP1_HABIN
ID HAP1_HABIN STANDARD; PRT; 1409 AA.
AC P44596;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adhesion and penetration protein precursor (EC 3.4.21.-).
GN HAP OR HI0248.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.C., Brandon R.C.,
RA Fine L.D., Fritchman D.L., Fuhmann J.L., Geoghegan N.S.M.,

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: PROBABLE PROTEASE, PROMOTES ADHERENCE AND INVASION BY
CC DIRECTLY BINDING TO A HOST CELL STRUCTURE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -1- SIMILARITY: Belongs to peptidase family S6.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A STOP CODON HAD TO
CC BE SKIPPED IN POSITION 710 TO PRODUCE THIS ORF.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32710; -; NOT_ANNOTATED_CDS.
CC MEROPS; S06.006; -.
DR TIGR; HI0248; -.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005545; Autotransporter.
DR InterPro; IPR000710; Peptidase_S6.
DR InterPro; IPR004899; Peptidase.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02395; IGA1; 1.
DR Pfam; PF03212; Peptidase; 1.
DR PRINTS; PR00921; IGASERPTASE.
DR TIGRfams; TIGR01414; autotrans_bar1; 1.
KW Hydrolyase, Serine protease, Transmembrane, Zymogen, Signal;
KW Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 ? POTENTIAL.
FT PROPEP ? ? ADHESION AND PENETRATION PROTEIN.
FT ACT SITE 250 250 HELPER PEPTIDE (POTENTIAL).
FT ACT SITE 250 250 BY SIMILARITY.
SQ SEQUENCE 1409 AA; 156797 MW; 63ABC937A84D16E CRC64;

Query Match 6.9%; Score 95.5; DB 1; Length 1409;
Best Local Similarity 23.9%; Pred. No. 1e+02;
Matches 53; Conservative 24; Mismatches 80; Indels 65; Gaps 9;

QY 64 NLTVLEVFNDYMPNSAVSKGELDANAMQHPYLEKD-----SQEGLNIVLVGN----- 115
DB 852 NLTLN-----NSTVTLNSAVSAS--SNVAPRRKSLTETPTTSAERFNTLVNGKLSG 905
QY 116 ---TFVYPLAGY-STKIKTINELKDGATVAVPNDPSNLARALILKEKGLIKDKNTNLF 171
DB 906 GTQFTGSSLFQYKSDKIKLSNDAGDYTLISV-----RNTGKEPTLRLQTLIESLDNPLS 961
QY 172 STTLDIYENPKKVIYKEVDSVAAARALDDVDVAVNNNVAQVGLTASGVVEVDKSP 231
DB 962 DKCKFTLEN-----DHVDAQALRKVIKNGEFLHNPI----- 995
QY 232 YVNIIVARADNKSKAIDPVKAYQD-----EVEAEAKKQ 267
DB 996 -----KEGELNDIVRAEQAQETLEAKQVQGTAKKQ 1026

RESULT 40
VATC_MOUSE
ID VATC_MOUSE STANDARD; PRT; 382 AA.
AC Q921G3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Vacuolar ATP synthase subunit C (EC 3.6.3.14) (V-ATPase C subunit)
 DE (Vacuolar proton pump C subunit).
 GN ATP6V1C1 OR ATP6C OR V-ATC.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Howell M.L.; Dean G.E.;
 RT "cDNA sequences for mouse vacuolar ATPase subunits."
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Subunit of the peripheral V1 complex of vacuolar ATPase.
 CC Subunit C is necessary for the assembly of the catalytic sector of
 CC the enzyme and is likely to have a specific function in its
 CC catalytic activity. V-ATPase is responsible for acidifying a
 CC variety of intracellular compartments in eukaryotic cells.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 CC H(+) (out).
 CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a
 CC peripheral catalytic V1 complex (components A to H) attached to
 CC an integral membrane V0 proton pore complex (components: a, c, c',
 CC c'', and d).
 CC -1- SIMILARITY: Belongs to the V-ATPase C subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U13839; AAC83084.1; -
 CC MGD: MG1:1913585; ATP6v1c1.
 DR GO: GO:0008553; F:hydrogen-exporting ATPase activity, phospho. . .; IDA.
 DR InterPro: IPR004907; V-ATPase_C.
 DR Pfam: PF03223; V-ATPase_C; 1.
 DR Hydrolase; ATP synthetase; Hydrogen ion transport.
 SQ SEQUENCE 382 AA; 43860 MW; CDF51D6EC9439 CRC64;

Query Match 6.9%; Score 95; DB 1; Length 382;
 Best Local Similarity 23.0%; Pred. No. 23; Mismatches 100; Indels 88; Gaps 16;
 Matches 70; Conservative 46;

QY	18	LAGCSNQSNEPAISKTAQTIKVG--VMAGPEQVAEVAQVAKKYNLTVELVEFND	74
DB	22	LHAATTKNNNLAVSSKFNIPDKVGLDVLVG-----LSDELAKLD-	62
QY	75	YAMPNSAVSKGELDANAMQHKFYLEKDSQKGLNVLIVG-----NTFYPLAGYSTK	127
DB	63	-AFVEGVKK-----VAQYMAVLEDSKQVQENLNASGVDLVTYITRFQMDAKYPIK	115
QY	128	--IKTLNELKDGATIAVPNDPSNLARALILEKQGLIKLKQNTNLF--TLLDIVENPKL	184
DB	116	QSLKNSSEIIAKGVQIDNDLKSASAYNNL--KGNLQNLERKNAGSLITRSIAE-----	168
QY	185	VIREVDTSVAARAIDVDVLAV-----NNVAGQVGLTAS-----ENGVFVEDKDSPTVN	234
DB	169	IYKQDFVLDSEYL--VTLLVVPKLNHNDMIKQYETLAEWVPRSSNVLSDDQSYLCN	226
QY	235	IIVARADNDSKAIDPF-----VYAGTDEVEAEA-----KKQFKDGV	272
DB	227	VTLFPR-----KAVDDPRKARENKFIVADFYNSEEMKADKEEMTELSTDKKQF--GP	278
QY	273	IKGW	276
DB	279	LVRW	282

Search completed: June 16, 2004, 11:07:10
 Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:04:11 / Search time 46 Seconds
(without alignments)
1893.109 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 1383
Sequence: 1 MNFGKINGICALASGIALAG.....TDEVEAKKQPKDGVIKGM 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp Unclassified: *
15: sp_rv1rus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	958.5	69.3	270	16	Q7VM93	Q7VM93 haemophilus
2	786.5	56.9	271	16	Q83MC6	Q83MC6 shigella fl
3	770	55.7	270	16	Q8CWC8	Q8CWC8 escherichia
4	752.5	54.4	269	16	Q8DFC1	Q8DFC1 vibrio vuln
5	739.5	49.6	269	16	Q87RS3	Q87RS3 vibrio para
6	685.5	43.3	272	16	Q8CVC2	Q8CVC2 escherichia
7	571.5	41.3	267	16	Q83F42	Q83F42 coxiella bu
8	571	41.3	259	2	O69441	O69441 legionella
9	563.5	40.7	266	16	Q8Y0X1	Q8Y0X1 raietonia s
10	560.5	40.5	278	16	Q8YD39	Q8YD39 bruceella me
11	554	40.1	262	16	Q8FV86	Q8FV86 bruceella su
12	554	40.0	262	16	Q7W317	Q7W317 bordetella
13	551	39.8	262	16	Q7WEW2	Q7WEW2 bordetella
14	527.5	38.1	265	16	Q7WFM1	Q7WFM1 bordetella
15	527.5	38.1	265	16	Q7WFM1	Q7WFM1 bordetella
16	527.5	38.1	265	16	Q7W4E3	Q7W4E3 bordetella

17	526.5	38.1	265	16	Q7VV70	Q7VV70 bordetella
18	524.5	37.9	258	16	Q921X5	Q921X5 rhizobium m
19	517.5	37.4	259	16	Q8XU08	Q8XU08 raietonia s
20	517	37.4	261	16	Q8RFN4	Q8RFN4 fusobacteri
21	517	37.4	261	16	Q86DA1	Q86DA1 rhizobium l
22	515.5	37.3	259	16	Q8U7G0	Q8U7G0 agrobacteri
23	513	37.1	307	16	Q89E28	Q89E28 bradyrhizob
24	509.5	36.8	286	16	Q9CIN7	Q9CIN7 lactococcus
25	507	36.7	259	16	Q8FXS4	Q8FXS4 bruceella su
26	507	36.7	261	16	Q8IEC4	Q8IEC4 bruceella me
27	506.5	36.6	261	16	Q88C15	Q88C15 pseudomonas
28	502.5	36.3	260	16	Q9HT68	Q9HT68 pseudomonas
29	501	36.2	269	16	Q8PGF0	Q8PGF0 xanthomonas
30	497.5	36.0	259	16	Q8XHS5	Q8XHS5 pseudomonas
31	497.5	36.0	266	16	Q8P4S9	Q8P4S9 xanthomonas
32	488	35.3	256	16	Q88RL7	Q88RL7 pseudomonas
33	487	35.2	257	16	Q87UN6	Q87UN6 pseudomonas
34	485	35.1	268	16	Q7V195	Q7V195 helicobacte
35	481	34.8	277	16	Q83YX8	Q83YX8 enterococcu
36	477.5	34.5	257	16	Q9PEE7	Q9PEE7 campylobact
37	469.5	33.9	273	16	Q8ELA4	Q8ELA4 oceanobacti
38	468	33.8	256	16	Q9PEE8	Q9PEE8 campylobact
39	467.5	33.8	271	16	Q26084	Q26084 helicobacte
40	464.5	33.6	286	16	Q9CIN6	Q9CIN6 lactococcus
41	463.5	33.5	273	16	Q92E25	Q92E25 listeria in
42	462.5	33.4	271	16	Q92U45	Q92U45 helicobacte
43	462.5	33.4	273	16	Q8YA74	Q8YA74 listeria mo
44	460	33.3	286	16	Q9CIN8	Q9CIN8 lactococcus
45	457.5	33.1	260	16	Q87UT0	Q87UT0 pseudomonas

ALIGNMENTS

RESULT 1

ID	Q7VM93	PRELIMINARY	FRT	270 AA.
AC	Q7VM93			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	28kDa outer membrane lipoprotein.			
GN	HLEA OR HD1100.			
OS	Haemophilus ducreyi.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			
OC	Pasteurellaceae; Haemophilus.			
OX	NCBI_TaxID=730;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=3500HP / ATCC 700724;			
RA	Munson R.S., Jr., Ray W.C., Manairas G., Sabo P., Mungur R.,			
RA	Johnson L., Nguyen D., Wang J., Forst C., Hood L.,			
RT	"The complete genome sequence of Haemophilus ducreyi."			
RL	Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, AF017153; AAP95966.1; -			
KM	lipoprotein; Complete proteome.			
SQ	SEQUENCE 270 AA; 29335 MW; FC077F10B13304AD CRC64;			

Query Match 69.3%; Score 958.5; DB 16; Length 270;
Best Local Similarity 71.5%; Pred. No. 1.4e-53;
Matches 188; Conservative 34; Mismatches 36; Indels 5; Gaps 2;

QY	14	SCIALAGSNGSNEPAAISKTAACITKGVWAGBQAAEYAGVAKKYLTVELVEFN	73
DB	13	STLALTGKPPANNQ---GDSASTIKKGVWAGBQAAEYAGVAKKYLTVELVEFN	68
QY	74	DYAMNSAVSKGELDANAMQKPYLEKDSQKGLNNTLVAGTFYPPLAGYSTIKITLNE	133
DB	69	DYALFNMAVSKGDDIDNMQKPYLDKDSQKGLNNTLVAGTFYPPLAGYSTIKITLNE	128
QY	134	LKDQATVAVPNDPSLALALLLEKQGLIKLKDNTNLFSTLTDIVENKYLIVKVDTSV	193
DB	129	LKSEAVVAVPNDPSNARALMLLEKQGLITLNDNTNLFATKLDIVDNKYLIVKVDTSV	188

QY 194 ARAIDVDVLA VNNNNYAGVGLTASENGVEVEDKDS PYNNITVARADNKSALDQDFYK 253
 DB 189 AAKLDDVDLA VNNNTTASQAGLSVNS -VFVEDKDS PYNNITVARADNKSALDQDFYK 247
 QY 254 AYQTEVEVEAEAKQFKDGVTKGW 276
 DB 248 AYQTEVEVEAEAKQFKDGVTKGW 270

RESULT 2

Q83MC6 PRELIMINARY; PRT; 271 AA.
 AC Q83MC6;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Putative lipoprotein.
 GN YAC OR SF0188 OR SF0190.
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 CX NCBI_Taxid=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=2272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang J., Zhang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.,
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 DR EMBL; AEO15055; AAN41850.1; -;
 DR EMBL; AEO16978; AAP15730.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; YaeC_
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRfams; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Lipoprotein; Complete proteome.
 SQ SEQUENCE 271 AA; 29417 MW; C3D7D01FB5CD2BF6 CRC64;

Query Match 56.9%; Score 786.5; DB 16; Length 271;
 Best Local Similarity 55.4%; Pred. No. 1.4e-42;
 Matches 133; Conservative 45; Mismatches 67; Indels 11; Gaps 3;
 QY 3 FKGINGICALASGIALAGCSNQNENPAISKTAAQTIKGVWAGPQAVAEVAGVAKX 62
 DB 5 FKTFAAVGALIGSLALVCGCGDEKDP-----NHIKGVIVGAEQVAVAGVAKX 56
 QY 63 YNLTVELVERNDYAMPNSAVSKGELDANAMQHPTLEKDSQKGLNNLVYGTFFVPLA 122
 DB 57 YGIDVELVTFNDYVLENEALSQKIDIDANAFQHKPYLDQQLKRGY-KLVAAGNTFVYPPIA 115
 QY 123 GYSTKIKTINELKDGAITAVPNDPSNLARALILEKQGLIKLKDNTNLFSTLIDVENPK 182
 DB 116 GYSKIKSLDELQDSQVAVPNDPTNLGSLILLKXVGLIKLKQVGLLPTVLIDVENPK 175

QY 183 KLVIKEVTSVAARAID--VDIAVNNNNYAGVGLTASENGVEVEDKDS PYNNITVARA 240
 DB 176 NKLTVELAPQLPSRLDQAIALAVINTTYSQIGLTPAKGIVFVEDKDS PYNNITVIRE 235
 QY 241 DNKDSKAIQDFYKAYQTEVEAEAKQFKDGVTKGW 276
 DB 236 DNKDAENYKFFQAYQSDVEVEAANKVFNAGVAKGW 271

RESULT 3

Q8CWC8 PRELIMINARY; PRT; 270 AA.
 AC Q8CWC8;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE D-mechonine-binding lipoprotein meq precursor.
 GN YAC OR C0238.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 CX NCBI_Taxid=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackert J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AEO16755; AAN78730.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR004478; YaeC_
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRfams; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Lipoprotein; Complete proteome.
 SQ SEQUENCE 270 AA; 29318 MW; 2EC4133C671EALBC CRC64;

Query Match 55.7%; Score 770; DB 16; Length 270;
 Best Local Similarity 55.1%; Pred. No. 1.6e-41;
 Matches 152; Conservative 45; Mismatches 67; Indels 12; Gaps 4;
 QY 3 FKGINGICALASGIALAGCSNQNENPAISKTAAQTIKGVWAGPQAVAEVAGVAKX 62
 DB 5 FKTFAAVGALIGSLALVCGCGDEKDP-----NHIKGVIVGAEQVAVAGVAKX 56
 QY 63 YNLTVELVERNDYAMPNSAVSKGELDANAMQHPTLEKDSQKGLNNLVYGTFFVPLA 122
 DB 57 YGIDVELVTFNDYVLENEALSQKIDIDANAFQHKPYLDQQLKRGY-KLVAAGNTFVYPPIA 115
 QY 123 GYSTKIKTINELKDGAITAVPNDPSNLARALILEKQGLIKLKDNTNLFSTLIDVENPK 182
 DB 116 GYSKIKSLDELQDSQVAVPNDPTNLGSLILLKXVGLIKLKQVGLLPTVLIDVENPK 174
 QY 183 KLVIKEVTSVAARAID--VDIAVNNNNYAGVGLTASENGVEVEDKDS PYNNITVARA 240
 DB 175 NKLTVELAPQLPSRLDQAIALAVINTTYSQIGLTPAKDGFVFDKSPYNNITVIRE 234
 QY 241 DNKDSKAIQDFYKAYQTEVEAEAKQFKDGVTKGW 276
 DB 235 DNKDAENYKFFQAYQSDVEVEAANKVFNAGVAKGW 270

RESULT 4

Q8DFC1 PRELIMINARY; PRT; 269 AA.
 AC Q8DFC1;
 DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE ABC-type metal ion transport system, periplasmic component.
 GN Vli0294.
 OS *Vibrio vulnificus*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; *Vibrio*.
 NCBI_TaxId=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMC6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RA "Complete genome sequence of *Vibrio vulnificus* CMC6";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB016797; AAC00827.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004437; Prok_Lipoprot_S.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Complete proteome.
 KW Complete proteome.
 SQ SEQUENCE 269 AA; 29234 MW; 774C7112BC3F1E15 CRC64;

 Query Match 54.4%; Score 752.5; DB 16; Length 269;
 Best Local Similarity 54.5%; Pred. No. 2e-40;
 Matches 146; Conservative 44; Mismatches 69; Indels 9; Gaps 2;

 QY 9 ICALSGIALAGCSNSNPAAISKTAQITIKYGVAGPEQAVAEVAGVAAEKXNLTV 68
 DB 11 VATAASALVLAGCDK-----EVDVNVKVGVIAGAEQVAEVAAKVAKEXNLTV 62
 QY 69 LVEFNDYVAMPNSAVSKGLDANAMQHPYLEKDSQKGLNNLVIGNTFVYPLAGYSTKI 128
 DB 63 LVTFDYVTPNALDDGSIDINAFQHKPYLDQVDRDY-KLTIGNTFVYPIAGYSKV 121
 QY 129 KTLNELKDGATIAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLTDIVENPKLVKE 188
 DB 122 KSVDELQGARIAVPNDPTNLGRSLLEKQGLITLSDVGLATFVDIVGNPKNITIE 181
 QY 189 VDTSVARAIDVDVLA VNNNYAGVGLTASENGVEFDKSPYNIIVARADNKDSKAI 248
 DB 182 LPAAGLPRLSDVDTLTIINTTYASSIDLSPEKDGVEFDKSPYNIIVAREVNAQNV 241
 QY 249 QDFKAYQTEVEAEAKKQFKDGVYKGM 276
 DB 242 QNFVKAQTEVEAEAKKIFKGVYKGM 269

 RESULT 5
 QSTRS3 PRELIMINARY; PRT; 269 AA.
 ID Q87RS3;
 AC Q87RS3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Lipoprotein Yaec.
 GN Vli0704.
 OS *Vibrio parahaemolyticus*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; *Vibrio*.
 NCBI_TaxId=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210653 / Serotype O3:K6;
 RA MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Nishimura M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shingawa H., Hattori M., Iida T.;
 RA "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
 RT distinct from that of *V. cholerae*.";
 RL Lancet 361:743-749 (2003).

DR EMBL; AB005075; BACS8967.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004437; Prok_Lipoprot_S.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Complete proteome.
 SQ SEQUENCE 269 AA; 29070 MW; 5D45BA921862ED8 CRC64;

 Query Match 53.5%; Score 739.5; DB 16; Length 269;
 Best Local Similarity 52.6%; Pred. No. 1.4e-39;
 Matches 141; Conservative 52; Mismatches 66; Indels 9; Gaps 2;

 QY 9 ICALSGIALAGCSNSNPAAISKTAQITIKYGVAGPEQAVAEVAGVAAEKXNLTV 68
 DB 11 VATAASALVLAGCDK-----EVDVNVKVGVIAGAEQVAEVAAKVAKEXNLTV 62
 QY 69 LVEFNDYVAMPNSAVSKGLDANAMQHPYLEKDSQKGLNNLVIGNTFVYPLAGYSTKI 128
 DB 63 LVTFDYVTPNALDDGSIDINAFQHKPYLDQVDRDY-KLTIGNTFVYPIAGYSKV 121
 QY 129 KTLNELKDGATIAVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLTDIVENPKLVKE 188
 DB 122 KSVDELQGARIAVPNDPTNLGRSLLEKQGLITLSDVGLATFVDIVGNPKNITIE 181
 QY 189 VDTSVARAIDVDVLA VNNNYAGVGLTASENGVEFDKSPYNIIVARADNKDSKAI 248
 DB 182 LEAPQLRSLDDVTLTIINTTYASSIDLSPEKDGVEFDKSPYNIIVAREVNAQNV 241
 QY 249 QDFKAYQTEVEAEAKKQFKDGVYKGM 276
 DB 242 QNFVKAQTEVEAEAKKIFKGVYKGM 269

 RESULT 6
 Q8CVK2 PRELIMINARY; PRT; 272 AA.
 ID Q8CVK2;
 AC Q8CVK2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Lipoprotein-28 precursor.
 GN NIBA OR C4583.
 OS *Escherichia coli* O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 NCBI_TaxId=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RA MEDLINE=22386234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Resch P.,
 RA Raeko D., Buckles E.L., Iliou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Domeneberg M.S., Blattner F.R.;
 RA "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic *Escherichia coli* O6".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 DR EMBL; AE016769; AAN83017.1; -;
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004437; Prok_Lipoprot_S.
 DR InterPro; IPR004478; Yaec.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Complete proteome.
 KW Complete proteome.
 SQ SEQUENCE 272 AA; 29414 MW; B0CFAD902B3D2F4 CRC64;

 Query Match 49.6%; Score 685.5; DB 16; Length 272;
 Best Local Similarity 49.8%; Pred. No. 3.9e-36;
 Matches 135; Conservative 54; Mismatches 71; Indels 11; Gaps 3;

 QY 8 ICALSGIALAGCSNSNPAAISKTAQITIKYGVAGPEQAVAEVAGVAAEKXNLTV 67
 DB 11 VATAASALVLAGCDK-----EVDVNVKVGVIAGAEQVAEVAAKVAKEXNLTV 62

[1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangenot S.,
 RA Arlat M., Billault A., Brotter P., Camus J.C., Catolico L.,
 RA Chandler C., Choisme N., Claudel-Renard C., Cumac S., Demange N.,
 RA Gaspin C., Lave M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sguier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weisenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646061, CAD14624.1, -;
 DR InterPro: IPR004872, Lipoprotein_9.
 DR InterPro: IPR006311, Tat.
 DR InterPro: IPR004478, YaeC.
 DR Pfam: PF03180, Lipoprotein_9; 1.
 DR TIGRFAMs: TIGR01409; TAT signal_seq; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 KM Complete proteome.
 SQ SEQUENCE 266 AA; 28515 MW; B718F52C9CFB236 CRC64;

Query Match 40.7%; Score 563.5; DB 16; Length 266;
 Best Local Similarity 43.2%; Pred. No. 2,3e-28;
 Matches 115; Conservative 47; Mismatches 87; Indels 17; Gaps 3;

QY 11 ALASGIALAGCSNQSNEPAISKTAAQTIKVGVMAPEQVAEVAQVAKKXNLTVELV 70
 DB ALAAGALA-----CGKPIKIGVAGPHAEIWEAVKVA-EGDGLKQIV 61
 QY 71 EFNDYAMPNSAVSKGELDANAMOHKPYLEKDSQKGLNVLIVGNTFYPIAGYSTIKIT 130
 DB EFNDYIOPRAALAGSLDANSYOHPIYDDQIATKRY-KFVSVGQITTFPWSVSKKKS 120
 QY 121 LKDLKDGARFGLFNDPTNGRALLLLQAGVILKKNAGFKASPRDVAENPKLFLVELD 180
 DB 121 LKDLKDGARFGLFNDPTNGRALLLLQAGVILKKNAGFKASPRDVAENPKLFLVELD 180
 QY 191 TSVAARATIDVDLAVNNNNYAGVGLTASENGVFVEDKDSPTVNTIIVAPADKDSKAIQ 250
 DB 181 AAQLPBSLDDLPAAVNGVYAKAGLDPRDIAIESPKGPANYIAVFAADKQDPWAK 240
 QY 251 FVKAYQTDVEAEAKKQFKDGVIKGW 276
 DB 241 LYKAVHSDAVKAFVTKYKDAVIVAM 266

RESULT 10
 Q8YD39 PRELIMINARY; PRT; 278 AA.
 AC Q8YD39;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN BMEI10338.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Brucellaceae; Brucella.
 CX NCB1_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delyecchio V.G., Kapactzi V., Redkar R.J., Patra G., Mujar C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonki L., Larsen N., D'Souza M., Bernat A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Lelesson J.-D.,
 RA Haeflorn R., Kyriades N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009671; AAL53580.1; -;

DR PIR: A13551; A13551.
 DR InterPro: IPR004872, Lipoprotein_9.
 DR InterPro: IPR006311, Tat.
 DR InterPro: IPR004478, YaeC.
 DR Pfam: PF03180, Lipoprotein_9; 1.
 DR TIGRFAMs: TIGR01409; TAT signal_seq; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 KM Complete proteome.
 SQ SEQUENCE 278 AA; 30606 MW; C0126189CC59B362 CRC64;

Query Match 40.5%; Score 560.5; DB 16; Length 278;
 Best Local Similarity 44.2%; Pred. No. 3,8e-28;
 Matches 118; Conservative 50; Mismatches 84; Indels 15; Gaps 5;

QY 11 ALASGIALAGCSNQSNEPAISKTAAQTIKVGVMAPEQVAEVAQVAKKXNLTVELV 70
 DB ALTVFA-----SAPSHAEKTKIKVINGEDEDVKVAEBGK-KHGINTERI 72
 QY 71 EFNDYAMPNSAVSKGELDANAMOHKPYLEKDSQKGLNVLIVGNTFYPIAGYSTIKIT 130
 DB TFNDYQNEALEKEIDEANAFQKPYIDQIKOHG-KISVAGTAVWPIGYSRKVK 131
 QY 131 LNELKDGATIVPNDPSNLARALILLEKQGLIKLKDNTNLFSTLDIVENPKLVIEVD 190
 DB 132 LELKDGATVGVFNDPTNEGRLRVLEBQGLIKLKPAGIATPIDIVENPKLVIEVD 191
 QY 191 TSVAARATIDVDLAVNNNNYAGVGLTASENGVFVEDKDSPTVNTIIVAPADKDSKAIQ 249
 DB 192 AGVGRSITDDLPALIVNNDAKAGL-KKEDATGESKKNPNVNTAVRTDLDQPMVK 250
 QY 250 DFVKAYQTDVEAEAKKQFKDGVIKGW 276
 DB 251 ALVASFQNDAVKALERAYKGTGIPAW 277

RESULT 11
 Q8FV86 PRELIMINARY; PRT; 278 AA.
 AC Q8FV86;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Lipoprotein, YaeC family.
 GN BRA0960.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Brucellaceae; Brucella.
 CX NCB1_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eissen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodele J.B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmiller S., Tetteilin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL: AE014589; AAN34131.1, -;
 DR PIR: A13551; A13551.
 DR TIGR: BR0960; -;
 DR InterPro: IPR004872, Lipoprotein_9.
 DR InterPro: IPR006311, Tat.
 DR InterPro: IPR004478, YaeC.
 DR Pfam: PF03180, Lipoprotein_9; 1.
 DR TIGRFAMs: TIGR01409; TAT signal_seq; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 KM Complete proteome.
 SQ SEQUENCE 278 AA; 30606 MW; C0126189CC59B362 CRC64;

Query Match 40.5%; Score 560.5; DB 16; Length 278;
 Best Local Similarity 44.2%; Pred. No. 3.8e-28;
 Matches 118; Conservative 50; Mismatches 84; Indels 15; Gaps 5;

QY 11 ALASGIALAGCSNOSNEPAISKTAQTIKVGWAGGEQAVAEKYNLTVEAV 70
 DB 25 ALTGFA-----SAPSHAEKTIKIGVIGGEDEDEWAKVABEGK-KGLINIERI 72
 QY 71 EENDVAMNSAVSKGELDANAMOHKPYLEKDSQKGLNLTIVGNTVYLAQYKIKT 130
 DB 73 TENDNPOPEALERELEIDANAFPHKPYDEQIKOHGY-KISVAGYIYVWPIGYSRKVK 131
 QY 131 INELKDGATIAVNDPSNLARALILEKQGLIKLKDNTNLFSTLIDVENPKLVKEVD 190
 DB 132 LEELKDGATVGVNDPTNEGRLALRYLEBQGLIKLKPDAGIATFIDIVENPKLKEIKED 191
 QY 191 TSVAARATIDVDLAVNNNVAQGVGLTASNGVVEKXD-SPYNNITVAADKDSKATQ 249
 DB 192 AGVVGKSIDDDALAVNDMAKAGL-KKEDATIGESKKNPNYNTFAVTDDLDQPMVK 250
 QY 250 DEKAYQTEVEAEAKKQFKDGVITKGV 276
 DB 251 ALVASFQNDVAVKAELERAYKGTGIPAW 277

RESULT 12
 ID Q7W317 PRELIMINARY; PRT; 262 AA.
 AC Q7W317;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP4046.
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_Taxid=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moutle S., Norbertzak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Bartell B.G., Maskell D.J.,
 RA "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL, BX640435; CAB39329.1; -
 KW Complete proteome.
 SQ SEQUENCE 262 AA; DAC18652307E892A CRC64;

Query Match 40.1%; Score 554; DB 16; Length 262;
 Best Local Similarity 42.5%; Pred. No. 9.2e-28;
 Matches 116; Conservative 55; Mismatches 86; Indels 16; Gaps 4;

QY 1 MNFGKINGICALAGSIALAGCSNOSNEPAISKTAQTIKVGWAGGEQAVAEKYNLTVEAV 60
 DB 3 MNFVR---SALLASAPVLGGAQ-----AEKLVGATQVPHAEILEV-KAL 47
 QY 61 EKNYLTVEAEFNDYAMPNSAVSKGELDANAMOHKPYLEKDSQKGLNLTIVGNTVEY 120
 DB 48 AKGVEALDIKVFITYVQPNQLADKQLDANFQHPYLDFTFNKDR-KTNLVSVGLVHVEP 106
 QY 121 LAGVSTKIKTNELKDGATIAVNDPSNLARALILEKQGLIKLKDNTNLFSTLIDVEN 180

DB 107 FGVSKKIKSLAEKDGATIAVNDPSNSGRALLLQKQGLIKLKDPNSNIVATPIDIAEN 166
 QY 181 PKLVIEVDTSVAARATIDVDLAVNNNVAQGVGLTASNGVVEKXDSPYNNITVAAD 240
 DB 167 PKLKFELERAMLEPSLDDLDLALINTVNALEAGVPTFDALFIEGADSPYNNITVAAR 226
 QY 241 DNKSKAIDPFVAYQTEVEAEAKKQFKDGVITKGV 273
 DB 227 DNKAPAVKLVNMLHSEAVRKFLIEKYGAVV 259

RESULT 13
 ID Q7WEW2 PRELIMINARY; PRT; 262 AA.
 AC Q7WEW2;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Putative exported protein.
 GN B4519.
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_Taxid=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moutle S., Norbertzak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Bartell B.G., Maskell D.J.,
 RA "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL, BX640450; CAB34882.1; -
 KW Complete proteome.
 SQ SEQUENCE 262 AA; 0DC80B1739FF4663 CRC64;

Query Match 40.0%; Score 553; DB 16; Length 262;
 Best Local Similarity 43.1%; Pred. No. 1.1e-27;
 Matches 116; Conservative 53; Mismatches 85; Indels 18; Gaps 5;

QY 1 MNFGKINGICALAGSIALAGCSNOSNEPAISKTAQTIKVGWAGGEQAVAEKYNLTVEAV 59
 DB 3 MNFVR---SALLASAPVLGGAQ-----AEKLVGATQVPHAEILEV-KPALA 48
 QY 60 KEKYNLTVEAEFNDYAMPNSAVSKGELDANAMOHKPYLEKDSQKGLNLTIVGNTVEY 119
 DB 49 KEGVEALDIKVFITYVQPNQLADKQLDANFQHPYLDFTFNKDR-KTNLVSVGLVHVE 105
 QY 120 PLAGVSTKIKTNELKDGATIAVNDPSNLARALILEKQGLIKLKDNTNLFSTLIDIVE 179
 DB 106 PFGVSKKIKSLAEKDGATIAVNDPSNSGRALLLQKQGLIKLKDPNSNIVATPIDIAE 165
 QY 180 NPKLVIEVDTSVAARATIDVDLAVNNNVAQGVGLTASNGVVEKXDSPYNNITVAAR 239
 DB 166 NPKLKFELERAMLEPSLDDLDLALINTVNALEAGVPTFDALFIEGADSPYNNITVAAR 225
 QY 240 ADNKSKAIDPFVAYQTEVEAEAKKQFKDGVITKGV 273
 DB 226 PDNKDAPAVKLVNMLHSEAVRKFLIEKYGAVV 259

RESULT 14
 ID Q7WSM1 PRELIMINARY; PRT; 262 AA.

AC Q7VSM1; 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP0385.
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 NCBI_Taxid=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama 1 / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett U.,
 Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 Bordetella parapertussis and Bordetella bronchiseptica."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640412; CAB44716.1; -
 KW Complete proteome.
 SQ SEQUENCE 262 AA; 28524 MW; E0A56B1A59FF4EC9 CRC64;

Query Match 39.8%; Score 551; DB 16; Length 262;
 Best Local Similarity 43.1%; Pred. No. 1.4e-27;
 Matches 118; Conservative 52; Mismatches 86; Indels 18; Gaps 5;

QY 1 MNFGKINGICALASGIALAGCSNQSNEPAISKTAQTIKVGVMGPEQAVAEVA-GQVA 59
 DB 3 MNFVR---SALLASATLALGGAQ-----AEKLVGATQVPAHEILEVAKPALA 48
 QY 60 KEKNTLVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVGNTFY 119
 DB 49 KGVVELDIKV--FTDYOQNLQIADKQIDANFQHPYIDTFNKOR-KNTLVSGLVHVE 105
 QY 120 PLAVGSKTKITNELDGATIAVPNDPSNLARALLILEKGLIKIKDNTLSTLLDIVE 179
 DB 106 PFGGISKTKISLAEKDGATIAIPNDPSNGRALLILQKGLKLDPSNIVATPIDIAE 165
 QY 180 NEPKLVKEVDTSVAAARADVDVAVNNVYAGQVLTASENGVEVEDSDSPYVNIIVAR 239
 DB 166 NEPKLKFRELAEAMLPSPFDLDLALINNVYALBGLVETRALPLEGADSYANLVAR 225
 QY 240 ADNKDSKATQDFYKAVQTEVEAEAKKQFQDGYI 273
 DB 226 PDKDAVAVKLVNLSAEVAKKFTIEKYKGAIV 259

RESULT 15
 Q7WFL1 PRELIMINARY; PRT; 265 AA.

AC Q7WFL1; 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Putative exported protein.
 GN B34169.
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 NCBI_Taxid=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-568;
 RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett U.,
 Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 Bordetella parapertussis and Bordetella bronchiseptica."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640449; CAB34533.1; -
 KW Complete proteome.
 SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B56685B CRC64;

Query Match 39.1%; Score 527.5; DB 16; Length 265;
 Best Local Similarity 39.6%; Pred. No. 4.6e-26;
 Matches 106; Conservative 57; Mismatches 94; Indels 11; Gaps 4;

QY 9 ICALASGIALAGCSNQSNEPAISKTAQTIKVGVMGPEQAVAEVAGVAKKYNLTVE 68
 DB 9 IAAFLGAAVRA-----QPALAD---KPKISVTNGPAAQIFEVVKQDA-AKQGLNTIQ 58
 QY 69 LVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNLTIVGNTFYVPLAGYSTKI 126
 DB 59 VIEFSDYQPVVALASGLDLSNSYQHPYLDNANADRGY-KLVGIATVTFPIGVYSKV 117
 QY 129 KTLNELDGATIAVPNDPSNLARALLILEKGLIKIKDNTLSTLLDIVEPKKLVYKE 188
 DB 118 KNLNELDGARIGIPNDPTNGRALLILQERGLIKLPEAGLAKATPIDIVENPKKLFIE 177
 QY 189 VDTSAARADVDVAVNNVYAGQVLTASENGVEVEDSDSPYVNIIVARADKDSKAI 246
 DB 178 LDAQQLRSLDDTASAVNTNPALEAGLDPSKDLVRESAESPYANLVVREGQKDAADL 237
 QY 249 QDFYKAVQTEVEAEAKKQFQDGYIKGM 276
 DB 238 RKLVSIVQSAVPVKFLLGYKGAIVAAW 265

RESULT 16
 Q7WAE3 PRELIMINARY; PRT; 265 AA.

AC Q7WAE3; 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP3723.
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 NCBI_Taxid=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-567;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett U.,
 Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 Bordetella parapertussis and Bordetella bronchiseptica."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640434; CAB39006.1; -

KM Complete proteome.
SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B566885B CRC64;

Query Match 38.1%; Score 527.5; DB 16; Length 265;

Best Local Similarity 39.6%; Pred. No. 4,6e-26; Indels 11; Gaps 4;

Matches 106; Conservative 57; Mismatches 94; Indels 11; Gaps 4;

QY 9 ICALASGIALAGCSNOSNEPAISKTAAQTIKYGVNAGPEQAVAEVAGVAKKXNLVVE 68

DB 9 LAAPALGAAYFA-----QPALAQD--KPLKIGVAGPHAQIFEVKQEA-AKQGLNIQ 58

QY 69 LVEFNDYAMPNSAVSNGEILDANAMQHKPYLEKDSQEKLNNTVYGNFTVYPLAGYSTKI 128

DB 59 VIEFSDYVQPNVALAGDLNANSYQHQPYLNNANRGY-KVLSIAKTVIFPIGVYSKV 117

QY 129 KLINEKDGATIAVPNDPSNLARALILIEKQGLIKLKDNTNLFSTLLDIVENPKLVIKE 188

DB 118 KLINEKDGARIGIPNDPTNGRALLLOEHGILKLRPEAGLKATPIDVENEKRLRFLTE 177

QY 189 VDTSVARAIDVDVLAVNNNNVAGVGLTASENGVEVEDSPYNNIYARADNDSKAI 248

DB 178 LDAQQLPRSLDDTDASAVNTNFMALFAGLDPSKDALVRESAESPYANVLVVRBQDKRADL 237

QY 249 QDFVKAQYOTDEVEAEAKKQPKDGIKGM 276

DB 238 KLVSVIYQSAFVREFFLGKTKGAVVAAM 265

RESULT 17

QYV770 PRELIMINARY; PRT; 265 AA.

ID QYV770; STRAIN=1021; PubMed=11481430;

DT 01-OCT-2003 (TREMblrel. 25, Created)

DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)

DE Putative exported protein.

GN BP2818.

OS Bordetella pertussis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=520;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;

RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achtman M., Atkin R., Baker S., Baaham D., Basson N., Chevreton I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,

RA Leather S., Mout R., Norbertczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitz B., Rutter S., Sanders M., Saunders D., Seeger K.,

RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,

RA Uwin L., Whitehead S., Barrell B.G., Maskell D.J.,

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica.",

RL Nat. Genet. 35:32-40(2003).

DR EMBL; BX640419; CAE43091.1; -

KM Complete proteome.

SQ SEQUENCE 265 AA; 28743 MW; 6B585474B5668844 CRC64;

Query Match 38.1%; Score 526.5; DB 16; Length 265;

Best Local Similarity 39.6%; Pred. No. 5,3e-26;

Matches 106; Conservative 56; Mismatches 95; Indels 11; Gaps 4;

QY 9 ICALASGIALAGCSNOSNEPAISKTAAQTIKYGVNAGPEQAVAEVAGVAKKXNLVVE 68

DB 9 LAAPALGAAYFA-----QPALAQD--KPLKIGVAGPHAQIFEVKQEA-AKQGLNIQ 58

QY 129 KLINEKDGATIAVPNDPSNLARALILIEKQGLIKLKDNTNLFSTLLDIVENPKLVIKE 188

DB 118 KLINEKDGARIGIPNDPTNGRALLLOEHGILKLRPEAGLKATPIDVENEKRLRFLTE 177

QY 189 VDTSVARAIDVDVLAVNNNNVAGVGLTASENGVEVEDSPYNNIYARADNDSKAI 248

DB 178 LDAQQLPRSLDDTDASAVNTNFMALFAGLDPSKDALVRESAESPYANVLVVRBQDKRADL 237

QY 249 QDFVKAQYOTDEVEAEAKKQPKDGIKGM 276

DB 238 KLVSVIYQSAFVREFFLGKTKGAVVAAM 265

RESULT 18

QYV770 PRELIMINARY; PRT; 258 AA.

ID QYV770; STRAIN=1021; PubMed=11481430;

DT 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE Putative outer membrane lipoprotein transmembrane.

GN R02888 OR SMC03157.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RX MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Boche G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaire V., Masuy D.,

RA Pohl T., Portetalle D., Puehler A., Purnelle B., Ransperger U.,

RA Renard C., Trebault P., Vandenbol M., Weidner S., Galibert F.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021.",

RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

DR EMBL; AL591792; CAC47467.1; -

DR InterPro; IPR004872; Lipoprotein_9.

DR InterPro; IPR004476; YaeC

DR Pfam; PF03180; Lipoprotein_9; 1.

DR TIGRFAWS; TIGR00363; TIGR00363; 1.

KM Complete proteome.

SQ SEQUENCE 258 AA; 27553 MW; 84917D53F26F435A CRC64;

Query Match 37.9%; Score 524.5; DB 16; Length 258;

Best Local Similarity 40.0%; Pred. No. 6,9e-26;

Matches 106; Conservative 57; Mismatches 83; Indels 19; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAAQTIKYGVNAGPEQAVAEVAGVAKKXNLVVE 71

DB 13 LAAGTAL-----ASTIKGVTFGEHAEIMEKVEVAARK-GHIDIEILE 54

QY 72 FNDYAMPNSAVSNGEILDANAMQHKPYLEKDSQEKLNNTVYGNFTVYPLAGYSTKI 131

DB 55 FSDYVVPQALADDLDNANSYQHQPYLNNANRGY-DIVSVGLITITTPGYSKSKSL 113

QY 132 NELKDGATIAVPNDPSNLARALILIEKQGLIKLKDNTNLFSTLLDIVENPKLVIKE 191

DB 114 DELDGATIAIPNDPTNGRALLLOEHGILKLRPEAGLKATPIDVENEKRLRFLTE 173

QY 192 SVARAIDVDVLAVNNNNVAGVGLTASENGVEVEDSPYNNIYARADNDSKAI 251

DB 174 AQLPRSLDDTDASAVNTNFMALFAGLDPSKDALVRESAESPYANVLVVRBQDKRADL 237

QY 252 VKAYOTDEVEAEAKKQPKDGIKGM 276

DB 234 VESYHDKVKAIFYEHKRGALIPSM 258

RESULT 19


```

08XU08
ID 08XU08 PRELIMINARY; PRT; 529 AA.
AC 08XU08;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-JUN-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative OUTERMEMBRANE signal peptide protein.
GN RSC3386 OR RS06100.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_Taxid=305;
RN [1]_SEQUENCE FROM N.A.
RC STRAIN=GM11000.
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoudat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlet M., Billault A., Brotlier P., Camus J.C., Cactolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cumac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646075; CAD16883.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR Pfam; PF03180; Lipoprotein_9; 2.
DR TIGRFAMs; TIGR00363; TIGR00363; 2.
KM Complete proteome.
SQ SEQUENCE 529 AA; 56737 MW; CD6D16319456D17 CRC64;

Query Match 37.4%; Score 517.5; DB 16; Length 529;
Best Local Similarity 39.1%; Pred. No. 56-25;
Matches 108; Conservative 65; Mismatches 84; Indels 19; Gaps 5;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAQTIKGVNAGPEQAVAEVAGVAKK 62
DB 265 FSRIRAAALPATAIAAGAHAEKSP-----IKLGTMSGDPAQIWEVAVQVAK-K 313
QY 63 YNLTVELVEFN DYAMPNSAVSKGELDANAMOHKPYLEKDSQEKGLNNVIYVNTFVYPLA 122
DB 314 DGLDRIKIIFENDYAOQNPALDAGDDPDANGFOHQPLDQGVAKGY-KIVNGLITVAEMG 372
QY 123 GYSTIKITINELDKGATIAVPNDPSNLRALILLEKGLIKLK-----DNTNLFSTTLDI 177
DB 373 FYAKKIKSLAALKEGKVGIIQNDPSNRRALLLLKAGVYIKLKAGAGNGSN--ATPRDV 430
QY 178 VENPKKIVKEVDTSVAARAIDVDLAVNNNYAGOVGLTASENVFVEDKDSPVNIIIV 237
DB 431 VENPKKIKIVELDSQQLPRSLDDLPAAALNTDYYAVKNGLTFPKDIALLEDRGFPYANLIA 490
QY 238 ARAADKDSKAIODPVKAYQTDVEVEAEAKKQFKDGYI 273
DB 491 VAEKDKNQPWVKTIVLHAYQSEVREKPIETQFKGALL 526

RESULT 20
Q8RFN4 PRELIMINARY; PRT; 261 AA.
AC 08RFN4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE ABC transporter substrate-binding protein.
GN FN0658.
OS Fuschacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_Taxid=76856;
RN [1]_SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;

```

```

RX MEDLINE=2186394; PubMed=11889109;
RA Kaparatel V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
RA Larsen N., D'Souza M., Mainas T., Pusch G., Haselkorn R.,
RA Forstner M., Kyrides N., Overbeek R.;
RA "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010577; AL94854.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR Pfam; PF03180; Lipoprotein_9; 1.
KM Complete proteome.
SQ SEQUENCE 261 AA; 28768 MW; 60AE194F467F0180 CRC64;

Query Match 37.4%; Score 517; DB 16; Length 261;
Best Local Similarity 41.3%; Pred. No. 21e-25;
Matches 112; Conservative 52; Mismatches 87; Indels 20; Gaps 4;

QY 3 FKGINGICALASGIALAGCSNOSNEPAISKTAQTIKGVNAGPEQAVAEVAGVAKK 62
DB 7 FGVYGAFLLSAG-ALMG-----TIKQATVPFAHILIKLKIPDLK-K 47
QY 63 YNLTVELVEFN DYAMPNSAVSKGELDANAMOHKPYLEKDSQEKGLNNVIYVNTFVYPLA 122
DB 48 QGVDLKIVEFTDYTPYLAALSDKEIDANFQHKPYLDKFIERRK-NLVSLGNVHVEPLG 106
QY 123 GYSTIKITINELDKGATIAVPNDPSNLRALILLEKGLIKLKDNNTLFSTTLDIYENPK 182
DB 107 LYSKIKISINDLKKGDITAIIPSPNSGRALILHNNKGVITLDPKPLFAETFDIYVNP 166
QY 183 KLVIKEVDTSVAARAIDVDLAVNNNYAGOVGLTASENVFVEDKDSPVNIIIVARADN 242
DB 167 KLKKRPKEVAGQLRILPDLVTAALINGVYALQANLSPKDSLIIEGKESPYANILIVAKGD 226
QY 243 KDSKAIDPVKAYQTDVEVEAEAKKQFKDGYI 273
DB 227 EKEDIOKLKALRSEKVKYINEXKSDGSV 257

RESULT 21
Q88DA1 PRELIMINARY; PRT; 284 AA.
AC 088DA1;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Outer membrane lipoprotein.
GN ML4784.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_Taxid=381;
RN [1]_SEQUENCE FROM N.A.
RC STRAIN=MAF0303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003005; BAB51370.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRFAMs; TIGR00363; TIGR00363; 1.
KM Lipoprotein; Complete proteome.
SQ SEQUENCE 284 AA; 30768 MW; 919CA8873A8C69BF CRC64;

```

Query Match 37.4%; Score 517; DB 16; Length 284;
Best Local Similarity 43.2%; Pred. No. 2.4e-25;
Matches 115; Conservative 48; Mismatches 85; Indels 18; Gaps 6;

14 SGIALAGCNSQNSNEPAISKTAAQ-----TIKVGWAGPEQVAEVAAGVAKKXN 64
20 SGIALAFAS-----AAVGSMAQHASLAEDKKAIKVIGLISEDDVRVVAQAAR-G 73
65 LTVLEVFENDYAMPNSAVSKGELDANAMQHPYLEKDSQEKLNVLVIGNTFFVPLAGY 124
74 LTVLEVFENDYTOPNPALENGEIDANAFQHPYLDNQIKTQGY-HIVRGYGVWPIGLY 132
125 STKIKTLNELKQATIAVNPDSNLARALILEKQGLIKLKNNTLFFSTLIDIVENPKKL 184
133 SKKYKAVADLPBQAVGVNPDSNEGRALRVQNEGVILKQGTGLATLADIAENPKKV 192
185 VIKEDVTSAAARIDVDVLAVVNNYAGVGLTASENGVEFVED-KDSPPVNTIYARADNK 243
193 EIKELDAGIVGSRVSEDLDAAVNTDALKSGLT-ENRIRAPRIADNPENFIAYKVGNE 251
244 DSKAIDDFKAVQOTDEVEAEAKKQK 269
252 NEAWVKTLVASTYQNDVAKAEFPDKVYK 277

RESULT 22
Q8U7G0 PRELIMINARY; PRT; 259 AA.

AC Q8U7G0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Outer membrane lipoprotein.
GN AtU4489 OR AGP L. 761.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Moo L.,
RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Boye D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RL Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Gueroio B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328 (2001).
DR EMBL; AE009377; AAL45283.1; -;
DR EMBL; AE008237; AAK8953.1; -;
DR PIR; AE3108; AE3108.
DR PIR; G98178; G98178.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.

DR TIGRFAMs: TIGR00363; TIGR00363; 1.
KW Lipoprotein: Complete proteome.
SQ SEQUENCE 259 AA; 27934 MW; F7D12ED13960861A CRC64;

Query Match 37.3%; Score 515.5; DB 16; Length 259;
Best Local Similarity 38.5%; Pred. No. 2.6e-25;
Matches 102; Conservative 58; Mismatches 94; Indels 11; Gaps 3;

12 LAGIALAGCNSQNSNEPAISKTAAQ-----LASTIKIGVPAHQIMEVKYKIAAK-GLDDITIE 55
6 LAASLALPTAQA-----LASTIKIGVPAHQIMEVKYKIAAK-GLDDITIE 55
72 FNDYAMPNSAVSKGELDANAMQHPYLEKDSQEKLNVLVIGNTFFVPLAGYSTKIKTL 131
56 FSDYVFNQALNDGELQANSPFQHPYLDNQIADRF-DIVSVGTITITPMGVSKKVKXNL 114
132 NELKQATIAVNPDSNLARALILEKQGLIKLKNNTLFFSTLIDIVENPKKLVKEVD 191
115 DELKDGATIGIIPDPPTNGSGLLVLSKGLVKEAGLVTPADITENPKNIQIVELDA 174
192 SVAAARIDVDVLAVVNNYAGVGLTASENGVEFVEDKDSPPVNTIYARADNDSKAIDPF 251
175 AQPLRSIDDTDAVINNTNVAITAGLNPKDAIMIESEKSPYANVIAVRADKDKPQVVKTL 234
252 VKAVQTDVEAEAKKQKPGVYKGV 276
235 VESYHSPEVKAFTLEKXNGTVPSPW 259

RESULT 23
Q89E28 PRELIMINARY; PRT; 307 AA.

AC Q89E28;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Outer membrane lipoprotein.
GN BL6903.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneke T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Matsumoto A., Idesawa K., Iriyuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197 (2002).
DR EMBL; AP005960; BAC52168.1; -;
DR InterPro; IPR004872; Lipoprotein_9.
DR Pfam; PF03180; Lipoprotein_9; 1.
KW Lipoprotein: Complete proteome.
SQ SEQUENCE 307 AA; 33948 MW; 8E4F3EC7C2CC8355 CRC64;

Query Match 37.1%; Score 513; DB 16; Length 307;
Best Local Similarity 41.1%; Pred. No. 4.7e-25;
Matches 102; Conservative 48; Mismatches 96; Indels 2; Gaps 2;

29 AAISKTAAQTIKVGWAGPEQVAEVAAGVAKKXNLTVLEVFENDYAMPNSAVSKGELD 88
62 AMLATATATIKIGVTPGPAQILAEVAKPIA-ARGLDIDQLLEFSYAYANALDAGEIQ 120
89 ANAMQHPYLEKDSQEKLNVLVIGNTFFVPLAGYSTKIKTLNELKQATIAVNPDSN 148
121 ANSFQNPYLDNQKADRGY-KIEAVGLTVFPIGVSKGKAKPADIPGQKVSIPNDPTN 179
149 LARALILEKQGLIKLKNNTLFFSTLIDIVENPKKLVKEVDTSVAAARIDVDVLAVVNN 208

Db 180 GGNVLLILDKGVKIKDGTGKPYVLDITENPKLKFIEVDAAQAAPRALDVAADAAINT 239
 QY 209 NVAGVGLTASNGVEVEDKDSFYVNIIVARADNCKSKAIQDFKAYQTFDEVEAEAKQOF 268
 Db 240 NVATQGLDVPKQPIILREDEKSPYVNLIAVRTADKQKPMVKLVDSYHPEVKEFVLTF 299
 QY 269 KQGVVIGM 276
 Db 300 KGAVLPSM 307

RESULT 24

Q9CIN7 PRELIMINARY; PRT; 286 AA.
 AC Q9CIN7
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Outer membrane lipoprotein precursor.
 GN PLPB OR IL0319
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=2135186; PubMed=11337471;
 RA Bolotin A., Winkler P., Manger S., Jallion O., Malarre K.,
 Weissendach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2003).
 DR EMBL; AF006269; AK004417.1; -.
 DR PIR; G86664; G86664.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR Lipoprotein; Complete proteome.
 SQ SEQUENCE 286 AA; 3163 MW; 0C3301DEFA422C0D9 CRC64;

Query Match 36.8%; Score 509.5; DB 16; Length 286;
 Best Local Similarity 40.5%; Pred. No. 7.2e-25;
 Matches 111; Conservative 58; Mismatches 90; Indels 15; Gaps 8;

QY 9 ICALASGIALAGCSNQNENPAISKTAQTIKYGVNAGPEQ--AAVEAVAGVAKKXKYL 66
 Db 17 IYAIATFIQF-QGKQAN-----KTVNKTIVKIGMTGKEDDSIMQVSKTAKDKYGIT 69
 QY 67 VELVERNDYAMNSAVSKELDANMOKHPYIEKDSQSEKGLNIVIGNTFVYPIAGYST 126
 Db 70 LKFTHTFTVYQNTALKNQDIDLNAFCHYAFI-KAMNKNANGNILVAIDTVISPIYSVK 128
 QY 127 KIKTINELKQATIVAPNDPSVLAALILLEKQGIKIKDNTNLFSTLIDIVENPKKVI 186
 Db 129 QLKNSIDIKEGTIAVPNDASNESALYVLSAGIILKIDVSGQTLATVKDITSNPKNVI 188
 QY 187 KEVDTSVAAARAIDVDVLAIVNNNNVAGVGLTASNGVEVE--DKDS-PYVNIIVARADN 243
 Db 189 KEIDASQTAFAALDSVDAVINNNVAVTAGLKSD-AIFTEPYNKSQGMINIIVANKDE 247
 QY 244 DSKAIQDFKAYQTFDEVEAEAKQKFD-GVTKGM 276
 Db 248 NNTVYKDVVAKAYEATKTKTIAKAYPDKSTIPAM 281

RESULT 25
 Q8FXS4 PRELIMINARY; PRT; 259 AA.
 AC Q8FXS4
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Lipoprotein, YaeC family.
 GN BR2175

OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Neilson K.E., Eisen J.A., Heidelberg J.F.,
 Read T.D., Dodson R.U., Umayam L., Brinkac L.M., Beanan M.J.,
 Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek O., Van Aken S.E.,
 Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AE014504; AANJ1064.1; -.
 DR TIGR; BR2175; -.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 KW Complete proteome.
 SQ SEQUENCE 259 AA; 27975 MW; A94FA42E3EBE74B CRC64;

Query Match 36.7%; Score 507; DB 16; Length 259;
 Best Local Similarity 40.0%; Pred. No. 9.1e-25;
 Matches 106; Conservative 51; Mismatches 98; Indels 10; Gaps 3;

QY 12 LASGIALAGCSNQNENPAISKTAQTIKYGVNAGPEQAAVEAVAGVAKKXKYLVE 71
 Db 5 LVAAITIALCTAGS-----AFAKTLRVGVTPGPHAQIMKRYEVAARK-GIEIDIOE 55
 QY 72 FNDAMPNSAVSKELDANMOKHPYIEKDSQSEKGLNIVIGNTFVYPIAGYSTIKTL 131
 Db 56 FSDVWNNMALADALEANCFQKHPYIDNQVADRGF-EIVSAVETVFPFAMYSRKLT 114
 QY 132 NELKQATIVAPNDPSVLAALILLEKQGIKIKDNTNLFSTLIDIVENPKKVIKVD 191
 Db 115 AELADGSTITAIPNDPTGGRALVLADNGILKIDSTGKLKVSADVATENRKNLKFVELD 174
 QY 192 SVAAARAIDVDVLAIVNNNNVAGVGLTASNGVEVEDKDSFYVNIIVARADNCKSKAIQDF 251
 Db 175 AQLPRESLEDVGAIVNTNVLLEGLDPRKADALVBEKAPKAYINILVVRTADKADAVKTL 234
 QY 252 VKAYQTFDEVEAEAKQKFDGVVIGM 276
 Db 235 IESYHSDEVDFIAKTFKGSVSPAM 259

RESULT 26

Q8YEC4 PRELIMINARY; PRT; 268 AA.
 AC Q8YEC4
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN BME11954.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=16W / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mufier C., Los T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Jablonowski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Detesseon J.-J.,
 Herscovici P., et al.;

"The genome sequence of the facultative intracellular pathogen *Brucella melitensis*."; *Proc. Natl. Acad. Sci. U.S.A.* 99:443-448(2002).

DR EMBL; AE009628; AAL53135.1; -.

DR PIR; AD3496; AD3496.

DR InterPro; IPR004872; Lipoprotein_9.

DR InterPro; IPR004478; YaeC.

DR Pfam; PF03180; Lipoprotein_9; 1.

DR TIGRfam; TIGR00363; TIGR00363; 1.

KW Complete proteome.

SEQUENCE 268 AA; 28948 MW; 1AF605B6B52E1F4B CRC64;

Query Match 36.7%; Score 507; DB 16; Length 268;
Best Local Similarity 40.0%; Pred. No. 9.5e-25;
Matches 106; Conservative 51; Mismatches 96; Indels 10; Gaps 3;

QY 12 LASGALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAE-VAGQVAKKYNLTVEL 71
DB 14 LVAATITLCTAGS-----APAKTLRVGVTGPPHAQIMKVEVAAKK-GIIEIIE 64

QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEGKLNVLIVGNTFYVPLAGSTKXITL 131
DB 65 FEDYVMPNVALADGALNANCFQHKPYLDNQVADRGF-EIVSVAETVNFPMAMYSRLKTI 123

QY 132 NEHDKGATINVPNDPSNIAKALILLEKQGLIKLKDNTNLFSTLIDIVENPKLVIREVT 191
DB 124 AELADGSTTALPNDPTNGRALLVLADNGLIKIDSTKGLKVSADVTEPKNKLFVELDA 183

QY 192 SVAAFAIDVDVLAVANNYAGQVGLTASENGVEFEDKDSRYVNIIVARADNKDSKALIDF 251
DB 184 AQLPRSLDVGAVINTNTYALAEGLDPRKADALVREKAPYITLIVRTADKADAVKTL 243

QY 252 VKAYQTDVEAEAKKQPFQDGYIKGW 276
DB 244 IESYHSDVKDFIAKTFKGSVSPAW 268

RESULT 27

088CLS PRELIMINARY; PRT; 261 AA.

AC 088CLS; 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Lipoprotein, putative.

GN PP5165.

OS Pseudomonas putida (strain KT2440).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=160488;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22423060; PubMed=12534463;

RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H., Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M., Brinkman L., Beaman M., Deboy R.T., Daugherty S., Kolonay J., Madupu R., Nelson W., White O., Peterson J., Knouri H., Hance I., Chris Lee P., Holtzapple E., Scanlan D., Tran K., Nozarez A., Utebbeck T., Rizzo M., Lee K., Kosack D., Moestl D., Medler H., Lander J., Sjeperand D., Hohnsbeil J., Straetz M., Helm S., Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tsemmler B., Fraser C.M.;

RT "Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440.";

RU Environ. Microbiol. 4:799-808(2002).

RL EMBL; AE016793; AAN0730.1; -.

DR TIGR; PP5165; -.

DR InterPro; IPR004872; Lipoprotein_9.

DR Pfam; PF03180; Lipoprotein_9; 1.

KW Complete proteome.

SEQUENCE 261 AA; 28159 MW; 9C1B3FE4687FEA13 CRC64;

Query Match 36.6%; Score 506.5; DB 16; Length 261;

Best Local Similarity 41.3%; Pred. No. 9.9e-25;
Matches 109; Conservative 52; Mismatches 88; Indels 15; Gaps 4;

QY 11 ALASGIALAGCSNOSNEPAISKTAQTIKVGWAGPEQVAE-VAGQVAKKYNLTVEL 69
DB 9 ALAALSFSGIA-----AAEKVLVATPVPHAEIIELELPTLAKSGVDLQIV 57

QY 70 VERNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEGKLNVLIVGNTFYVPLAGSTKXITL 129
DB 58 --FTDYVQPNVQDQKLDNANFYQLPYQNFEGGTHLETVG-VHPEFGGYSKVK 114

QY 130 TNELKDGATINVPNDPSNIAKALILLEKQGLIKLKDNTNLFSTLIDIVENPKLVIREV 189
DB 115 ALSLEKGAITVALPNEGSSNGRALLLLQKRGITLWDPRKALATPDIADENPKKXPRFL 174

QY 190 DTSVAFAIDVDVLAVANNYAGQVGLTASENGVEFEDKDSRYVNIIVARADNKSKALIQ 249
DB 175 ESANLPVQLQVDLMDINTNTYALAEGLNPKADLVLEGADSPYVFLVARPNKSEALIQ 234

QY 250 DFVAYQTDVEAEAKKQPFQDGYI 273
DB 235 KLVALTSPVKAFAKKGAVL 258

RESULT 28

09HT68 PRELIMINARY; PRT; 260 AA.

AC 09HT68; 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Probable TonB-dependent receptor.

GN PA5505.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 15692 / PAO1; MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Watterer P., Hickey M.J., Brinkman F.S.L., Huinagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Gentry L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Lattig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saiter M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen.";

RU Nature 406:959-964(2000).

RX EMBL; AE004963; AAG08890.1; -.

DR PIR; E82957; E82957.

DR CO; GO:0016020; C:membrane; IEA.

DR CO; GO:0004872; F:receptor activity; IEA.

DR CO; GO:0005315; F:transporter activity; IEA.

DR CO; GO:000610; F:transport; IEA.

DR InterPro; IPR004872; Lipoprotein_9.

DR InterPro; IPR000531; TonB_box.

DR Pfam; PF03180; Lipoprotein_9; 1.

DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.

KW Receptor; Complete proteome.

SEQUENCE 260 AA; 28065 MW; 1E553A0D5C9BEA1A CRC64;

Query Match 36.3%; Score 502.5; DB 16; Length 260;
Best Local Similarity 42.6%; Pred. No. 1.8e-24;
Matches 106; Conservative 52; Mismatches 84; Indels 7; Gaps 3;

QY 29 AAISKTAQTIKVGWAGPEQVAE-VAGQVAKKYNLTVELFNDYAMPNSAVSK 84
DB 12 AALGLTAAQAASITVAATPVPNHAELINVPKPLAKSGVDLKIK--FTDYVQPNVQVSE 69

QY 85 GELDANAMQHKPYLEKDSOEGKLNVLIVGNTFYVPLAGSTKXITLNEHDKGATINVPN 144

Db 70 KRLDANFQHQPYLDEFNKAG-TDLVAVNGVHIEPLGAVSSKKYKLDLPSGATVIVN 128
 QY 145 DPNLNARALILEKQGLIKLKONTNLFSTLTDIVENPKLVKEVDTVAARAIIDVDIA 204
 Db 129 DATTNGRRALILLIDKAGVITKDKNSITATPKDIVDPKNIKIRELEAATLPRLTQVDYA 188
 QY 205 VVANNYAGVGLTASENGVEVEDKSPYVNIIVARADNDSKAIODPVAYOTDEVEEA 264
 Db 189 LINTNVALEKKNPTDATALIEGSDSPYVNIIVARPDNDSDMQLALALHSAETKOPT 248
 QY 265 KQKQKQGVIT 273
 Db 249 QEKYKQAVV 257

RESULT 29

Q8PFG0 PRELIMINARY; PRT; 269 AA.
 AC Q8PFG0;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Outer membrane protein.
 GN XAC3667.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=2022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Camavan P., Cardozo J., Chambergo F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cristina-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.U.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Fournghieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Martins E.C., Machado M.A., Madela A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Mendez C.F.M., Miyaki C.Y., Moon D.H.,
 RA Pereira H.A., Rossi A., Sena U.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.,
 RA "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities".
 RL Nature 417:459-463(2002).
 RL EMBL: AE012017; AM38510.1; --
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0005125; F:transporter activity; IEA.
 DR GO: GO:0006810; F:transport; IEA.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_Lipoprotein_S.
 DR InterPro: IPR000531; TonB_DoxC.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 DR PROSITE: PS00013; PROKAR LIPOPROTEIN; 1.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 269 AA; 29097 MW; 75F2ACEB07C2B806A CRC64;

Query Match 36.2%; Score 501; DB 16; Length 269;
 Best Local Similarity 41.6%; Pred. No. 2.3e-24;

Matches 109; Conservative 47; Mismatches 98; Indels 8; Gaps 4;
 QY 13 ASGIALAGSNGSNBPAAISKTAOTIKYGVWAGPEQAVAEQ-VAKKYNLTVELVE 71
 Db 12 AATLALAGCGSSG-----SGGGDTLVVATVAVHAETLEVEPLKQGVKLDVAV- 65
 QY 72 ENDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNVLVVGNTFVYPLAGYSTIKITL 131

Db 66 FNDYVQPDQVQKQIDVNYVQTEPEPYLDVAVNRD-KSQLVTVGVHIEPFGAVSRFRAL 124
 QY 113 NELKQATIAVNPDPNSLNARALILEKQGLIKLKONTNLFSTLTDIVENPKLVKEVDT 191
 Db 125 AELPTGAVVAVPNDPSNNSRALLIDKAGVITKDKNDPSNALSQRDIVENPKLVKEVDT 184
 QY 192 SVARAIIDVDLAVVNNYAGVGLTASENGVEVEDKSPYVNIIVARADNDSKAIODF 251
 Db 185 AMLPRIVDQVDLALINTVYALDAGLNPTDATALIESKDSPPYVNIIVARADNDSKAIODF 244
 QY 252 VKAVOTDEVEAEAKQFQDGVIT 273
 Db 245 AKALTSPVYKAFIQKTKGAVL 266

RESULT 30

Q9HX85 PRELIMINARY; PRT; 259 AA.
 AC Q9HX85;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein PA3931.
 GN PA3931.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.A., Coulter S.N., Folger K.R., Kae A., Larys K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Slier M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen".
 RL Nature 406:959-964(2000).
 RL EMBL: AE004811; AA007318.1; --
 DR PIR: G83153; G83153.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR Pfam: PF03180; Lipoprotein_9; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 259 AA; 28101 MW; C5DF8C298BD233BA CRC64;

Query Match 36.0%; Score 497.5; DB 16; Length 259;
 Best Local Similarity 41.4%; Pred. No. 3.7e-24;
 Matches 110; Conservative 51; Mismatches 88; Indels 17; Gaps 4;

QY 9 ICLASIALAGSNGSNBPAAISKTAOTIKYGVWAGPEQAVAE-VAGVAKKYNLTVE 67
 Db 7 LTLASVT-----SVSASAEXLIVAAATPIPAEILIEDIKPTLAKEGVDTQI 53
 QY 68 ELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNVLVVGNTFVYPLAGYSTK 127
 Db 54 KV--FDYVQPNQVAKERLDANYPQTLPLEFNFGKGNLVTVVG-VAVEFQGYSKR 110
 QY 128 IKTUNLEIKQATIAVNPDPNSLNARALILEKQGLIKLKONTNLFSTLTDIVENPKLVIT 187
 Db 111 YKSLAEHPDQATVAIPNEGNSGRALLILQKGLIKLKONNALATPKDIAENPKLVIT 170
 QY 188 EVDTVAARAIIDVDLAVVNNYAGVGLTASENGVEVEDKSPYVNIIVARADNDSKAIODF 247
 Db 171 ELSEALIPRLVDVDDLINTVYALEKKNPAKDALVLEDRSDPYVNIIVARADNDSKAIODF 230
 QY 248 IQDFVAYQTDVEAEAKQFQDGVIT 273
 Db 231 IKTLSAALTSPVYKAFIEKKYAGAVV 256

DT 01-JUN-2003 (Tremblrel. 24, last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
 DE D-mechonine-binding lipoprotein MetQ.
 GN METQ-2 OR PSPT05260.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feldblum T., Gwin M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Dougherty S.,
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
 RA White O., Fraser C., Collier A.,
 RA "Complete sequence of Pseudomonas syringae."
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AB016875; AA058686.1; -
 DR TIGR; PSPT05260; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:000610; F:transport; IEA.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000531; TonB boxC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00430; TONB DEPENDENT_REC_1; 1.
 DR Lipoprotein; Complete proteome.
 KM SEQUENCE 257 AA; 27806 MW; 856EE817AB9E10E8 CRC64;
 SQ
 Query Match 35.1%; Score 487; DB 16; Length 257;
 Best Local Similarity 42.9%; Pred. No. 1.7e-23;
 Matches 105; Conservative 45; Mismatches 91; Indels 4; Gaps 3;
 QY 30 AISKTAQITIKYVAGMPGQAAE-VAGOVAKEKNLTVELFENDYAMPNSAVSKGELD 88
 DB 13 AISAQNNETLTVAASVAPHAETLEFVKPLAKEGVDELTKV--FNDYIOPNAQVQSKRID 70
 QY 89 ANAMQHKPYLEKDSQKGLNIVIGNTFVYPLAGSTKIKITLNEKOGATAVENDPSN 148
 DB 72 ANPFGQPLDDENNKGG-TDLVAAKAVPEFGAVSDKFKLDLPGANALNDATN 129
 QY 149 LRAALILEKQGLIKKNTNTLFTTLDIVENPKKLVKEVDTSVAARAIDVDLAANN 208
 DB 130 EGRALLLLAKAGILITKDPNSNLSKPSDVNNPKDLKPELEAACTLPVLTQVDLALINT 189
 QY 209 NYAGQVGLTASENGVEFEDKSPYVNIIVARADNDSKAIQDFVAKYOTDEVEAKKQF 268
 DB 190 NYALSKLDPTKCALIIEGADSPYANILITRPDKDSDAIKCLVALDLOSEVKTFLAEKY 249
 QY 269 KQGVV 273
 DB 250 KQAVV 254
 RESULT 34
 QY7195 PRELIMINARY; PRT; 268 AA.
 AC QY7195;
 DT 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
 DE Hypothetical protein yaeC.
 GN YAE C OR HH0713.
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=32025;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 51449 / 3B1;
 MEDLINE=22709201; PubMed=12810954;
 RX

RA Suerbaum S., Josephans C., Sterzenbach T., Drescher B., Brandt P.,
 RA Bell M., Droege M., Fatmann B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
 RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.,
 RA "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus."
 RT Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
 RL EMBL; AB017146; AAP77310.1; -
 DR Hypothetical protein; Complete proteome.
 KM SEQUENCE 268 AA; 29314 MW; 20948E0B2918109F CRC64;
 SQ
 Query Match 35.1%; Score 485; DB 16; Length 268;
 Best Local Similarity 40.0%; Pred. No. 2.4e-23;
 Matches 104; Conservative 56; Mismatches 94; Indels 6; Gaps 5;
 QY 15 GIALAGCSNNGNEPAISKTAQTI-KVGVVAGPEQAAVEVAGVAKENLTVELVEFN 73
 DB 11 GLVEFGCSDEKN--TSQGTAEKVLKVGATPVPHAETLEFI-KPDLKEGIDIDQVQFT 67
 QY 74 DYAMPNSAVSKGELDANAMQHKPYLEKDSQKGLNIVIGNTFVYPLAGSTKIKITLNE 133
 DB 68 DVTFNVSANDGSLDANHOKPFLDLAKDKRG-NLESASIHIEPLGFYSHKFKIDS 126
 QY 134 LKDGATIAVPNDPSNLARALILEKQGLIKKNTNTLFTTLDIVENPKKLVKEVDTSV 193
 DB 127 LPQGTIAIPNDPSNGRALLLDSKGIKLDATNNLNATLSDIVENPKKIKIPVEAAL 186
 QY 194 AARAIDVDLAANNVAGVGLTASENGVEFEDKSPYVNIIVARADNDSKAIQDFVK 253
 DB 187 LPRLLDDVDGAVINGNALQGLKSD-ALFLESGQSPYANILITVQSTRVADENILQTLKN 245
 QY 254 AYOTDEVEAKKQFQGVV 273
 DB 246 ALQSQKXKDFIQHYQGEIV 265
 RESULT 35
 QY832Y8 PRELIMINARY; PRT; 277 AA.
 AC QY832Y8;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
 DE Lipoprotein, yaeC family.
 GN EF2080.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=151;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
 RA Dougherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Khouri H.,
 RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.,
 RA "Role of mobile DNA in the evolution of vancomycin-resistant
 RT Enterococcus faecalis."
 RT Science 299:2071-2074(2003).
 RL EMBL; AB016953; AA081813.1; -
 DR TIGR; EF2080; -
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR000437; Prok_lipoprot_5.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 DR Lipoprotein; Complete proteome.
 KM SEQUENCE 277 AA; 30445 MW; 0A64B63154B79551 CRC64;
 SQ
 Query Match 34.8%; Score 481; DB 16; Length 277;
 Best Local Similarity 36.8%; Pred. No. 4.5e-23;
 Matches 102; Conservative 62; Mismatches 101; Indels 12; Gaps 5;

```

QY 3 FKGINGICLALAGSIALAGCSNSENPAISKTAAQITKVGWAGPEQAVAEVAGVAK 62
DB 4 FSKTILILVLA-FTTAGASGS-----VKDTKTEIVKLGVGKRNDEWESYKRLKK- 56
QY 63 YNLTVELFENDPAMPNSAVSKGELDANAMQHPYLEKDSQEKGINLTVIGNTFVYPLA 122
DB 57 -NIDLOVERDYTOGNALAEKEIDLNAPFOHOIFLDNNKNGH- TKLVISINTVAPLG 114
QY 123 GYSTKIKTINELKDGATIAVPNDPSNLARALILKEQGLIKDKNTNLFSTLLDVENPK 182
DB 115 IYANKKIDITIKIDGGEIAPNDPTNGRALILQTAGIKVDPAKOQLPTVSDITENKR 174
QY 183 KVIKKEVDTSVARRAIDVDLAVNNNNYAGQGLTASENGVEVE--DKSPYNNIIVAR 239
DB 175 QLTITLDTQTNALQDVPDASVINGMAVDAGYTDKPAIPIEPVNEKAKRYNIVAR 234
QY 240 ADNKDSKALQDFVKAQYQDVEAEAKKQFKGVYIKGW 276
DB 235 EEDQENKLVQKVEEYQOEETKKVIAETSKGANVPM 271

RESULT 36
Q9PEP7 PRELIMINARY; PRT; 257 AA.
ID Q9PEP7
AC Q9PEP7
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Putative periplasmic protein.
GN C00772C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
CX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RA MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jorgels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL; AL139076; CAB73037.1; -.
DR PIR; D81348; D81348.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000531; TonB box.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
DR Complete proteome.
SQ SEQUENCE 257 AA; 28638 MW; D03782C849297381 CRC64;

Query Match 34.5%; Score 477.5; DB 16; Length 257;
Best Local Similarity 41.5%; Pred. No. 6,9e-23;
Matches 100; Conservative 50; Mismatches 78; Indels 13; Gaps 5;

QY 30 AISKTAQITKVGWAGPEQAVAEVAGVAK--EKNTLVELFENDPAMPNSAVSKGEL 87
DB 15 SINNALETTTVAATPVPH--AEILEQVPRDEKQGLKEIKETFDYVLEPNLAVNGSA 71
QY 88 DANAMQHPYLEKDSQEKGINLTVIGNTFVYPLAGSTKIKTINELKDGATIAVNDPS 147
DB 72 DANFQITVLEENKNGK-TKLKVAALHIEPMAYSKYKSLDDIKGVKAIANDPT 130
QY 148 NLARALILKEQGLIKDKNTNLFSTLLDVENPKLVKEVDTSVARRAIDVDLAVN 207

```

```

DB 131 NESALDIIAKKGLVKKDKA--LKTPLDIIIDPKKIKFEVLEKPAQLPALNDVDPVAVIN 188
QY 208 NNTVAGVGLTASENGVEVEKDSPPYNNIIVARNDNSKALQDFVKAQYQDVEAEAKKQ 267
DB 189 SNVALSNLNPADDSVFTEDEKSPYANILVVRGHENDPKIKALIQHLSDKL-----KQ 243
QY 268 F 268
DB 244 F 244

RESULT 37
Q9BEL4 PRELIMINARY; PRT; 273 AA.
ID Q9BEL4
AC Q9BEL4
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical conserved protein.
GN O33327.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
CX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RA MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments."
RT Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004604; BAC15283.1; -.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok Lipoprot_5.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 273 AA; 29904 MW; B05CE0F7A78B211F CRC64;

Query Match 33.9%; Score 469.5; DB 16; Length 273;
Best Local Similarity 34.8%; Pred. No. 2,4e-22;
Matches 96; Conservative 66; Mismatches 111; Indels 3; Gaps 3;

QY 1 MFKINGICLALAGSIALAGCSNSENPAISKTAAQITKVGWAGPEQAVAEVAGVAK 60
DB 1 MKKGLTAIVFLAVL-LAAGSSESSGASEGLYDEKLVIGTAGHHEILEKVTLEAE 59
QY 61 EKNLTVELFENDPAMPNSAVSKGELDANAMQHPYLEKDSQEKGINLTVIGNTFVYPL 120
DB 60 EE-GINIEVEFTDYVWPNVALDGEGLDNIPOTEPFNMIREDRNL-DIVKSFDTVTFP 117
QY 121 LAGSTKIKTINELKDGATIAVPNDPSNLARALILKEQGLIKDKNTNLFSTLLDVEN 180
DB 118 MGIYSLDVGVSELPREGATTGLPDPNLSGRALLPFOAGLIELDPETGINSIVRDIKYN 177
QY 181 PKLVKKEVDTSVARRAIDVDLAVNNNNYAGQGLTASENGVEVEKDSPPYNNIIVARA 240
DB 178 NGNIEFVFLDSAQARQLEIEDAALNTNPAILEGFTPTDPAIPIEPVNEKAKRYNIVARA 237
QY 241 DNKDSKALQDFVKAQYQDVEAEAKKQFKGVYIKGW 276
DB 238 ENEDDEITQKLADIYRSEEVKTFIEEFSGSLVPSW 273

RESULT 38
Q9PEP8 PRELIMINARY; PRT; 256 AA.
ID Q9PEP8
AC Q9PEP8
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

```


DE Putative periplasmic protein.
GN C00771C.
RA Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holt R.D.,
RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.,
RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000)
RU EMBL: AL139076; CAB73036.1; -
DR PIR: C81348; C81348.
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam: PF03180; Lipoprotein_9; 1.
KW Complete proteome.
SQ SEQUENCE 256 AA; 28843 MW; E64BA955ED452470 CRC64;
Query Match 33.8%; Score 468; DB 16; Length 256;
Best Local Similarity 42.2%; Pred. No. 2,8e-22;
Matches 103; Conservative 52; Mismatches 83; Indels 6; Gaps 4;
QY 31 ISKTAQTKVGVAGPEQVAEAGVAKKX-YNLTVELVENDYAMPNSAVSKGELDA 89
DB 15 LSLFLQNTTIGTATPFGSLHMDPRKNGYEL--KIVERSDYILNRLAEKELDA 72
QY 90 NAMQHPYLEKDSQSEKGLNNLVIVNTFVYPLAGYSTKIKTLNELKGATIAVPNDPSNL 149
DB 73 NLYQHPFPLEEYVLLKKG-SNLVATTPVLLAPGVYSKIKLENLEKGAVALIPNDATNE 131
QY 150 ARALLILEKQGLIKKDNNTLFTLDIYENPKKVIKEVDNSVAAARAIDVDVLAIVNNN 209
DB 132 SRLLELEKAKLEL--NKNTLKTPIIDINKPKKLEFILLKAAQLFRALDDVDIALINSN 189
QY 210 YAOQVGLTASNGFVEDKDSPYVNIIVARADNKSKAIDQFVKAQOTDEVEAEAKKQPK 269
DB 190 FALGAGLNSKDTIFEDKNSPYVNVVVRSEKNSKTKVIDELIRSKFAIINEHYK 249
QY 270 DGYI 273
DB 250 DILI 253
RESULT 39
Q26084 PRELIMINARY; PRT; 271 AA.
ID C26084;
AC C26084;
DT 01-JUN-1998 (TrEMBLrel. 05, Created)
DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Outer membrane protein.
GN HP1564.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush D., Zhou L., Kirkness S.F., Peterson S.,
RA Loftus B., Richardson J., Dodson R., Khalak H.G., Gocke A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Goeysne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Mallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
DR EMBL: AE000654; AAD08604.1; -
DR PIR: D64715; D64715.
DR TIGR: HP1564; -
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam: PF03180; Lipoprotein_9; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 271 AA; 30151 MW; 23786AEFB5C6CSB CRC64;
Query Match 33.8%; Score 467.5; DB 16; Length 271;
Best Local Similarity 42.1%; Pred. No. 3.2e-22;
Matches 99; Conservative 46; Mismatches 85; Indels 5; Gaps 3;
QY 39 IKGVAGPEQVAEAGVAKKX-YNLTVELVENDYAMPNSAVSKGELDANAMQHPYL 98
DB 39 LKGVANPVFAQIIQSVVDLKKK-GIKVLVSFTDYVLPALINDSLDANVFOHRPYL 97
QY 99 EKDSQKGLNNLVIVNTFVYPLAGYSTKIKTLNELKGATIAVPNDPSNLARALLLEK 158
DB 98 DRFNDRKM-HLVGLANIHVAPLRFYSQKTTIDIKLKKGSYIAVPNDPANOGRLILHLK 156
QY 159 QGLIKLKDNTLFTLDIYENPKKVIKEVDNSVAAARAIDVDVLAIVNNNVAQVGLTA 218
DB 157 QGLIALKQPSNLVATFPIVGNPNYIKIKPLEALLPKVLGVDGALITGNYVALQAKLTG 216
QY 219 SENGVFVEDKDSPYVNIIVARADNKSKAIDQFVKAQOTDEVEAEAKKQPKDGYI 273
DB 217 A--LFSDKDSFYANLIAREDNAQDEALKTLEALQSEKTKRFIDTYGAIL 268
RESULT 40
Q26084 PRELIMINARY; PRT; 286 AA.
ID Q26084;
AC Q26084;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Outer membrane lipoprotein precursor.
GN PLPC OR IL0320.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauer S., Jallion O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL: AE006269; AAK04418.1; -
DR PIR: H86664; H86664.
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam: PF03180; Lipoprotein_9; 1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 286 AA; 31554 MW; 398725DC477A4D65 CRC64;
Query Match 33.6%; Score 464.5; DB 16; Length 286;
Best Local Similarity 40.5%; Pred. No. 5.3e-22;
Matches 113; Conservative 54; Mismatches 101; Indels 11; Gaps 7;
QY 1 MNFKINGICALASGIALAGCSNOS-NEPAISKTAQTKVGVAG--EQAVAEVAGQ 57
DB 1 MNPKNRNIVAVIYAVLAVAFSLNHQASKSKTAKTVAGVIGSGDKODQEWKSVAK 60
QY 58 VAKKYNLTVELVENDYAMPNSAVSKGELDANAMQHPYLEKDSQKGLNNLVIVNTF 117

Db 61 TAKEKYDLKLFVYFSPDYNOPEALLSGDIDVNAFOSYNYV-KTMNKAHKSIDIYAVGNTY 119
QY 118 VYPLAGYSTKIKITLNEIKOGATIAVPNDPSNLARALILEKOGILKL--KDNTNLFSTTL 175
Db 120 ITPWHIYSKEVSKLSDLKBSGTVAIIPDASNEBRALFVLOSAGLILKLTSDSKLVGLP- 178
QY 176 DIVENPKGLVYKEVDTSVAAARAIIDVDLAVNNNNYAGQVGLTASENGVE--DKDSPY 232
Db 179 DITENPHQLKFKVEYDASQTPRALDSVALSVVNTNYATAASLPNSES-VYMEPLNKTSAGY 237
QY 233 VNIIVAAADNKSKAIODFYKAYQTDVEAEAKKQFKDG 271
Db 238 INFIAATSKKEKNKKNVYKEVAKAYASKATEKAIKEQFPDG 276

Search completed: June 16, 2004, 11:08:08
Job time : 49 secs

Tue Jun 22 11:04:04 2004

us-10-018-672-2.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:05:36 ; Search time 22 Seconds
(without alignments)
647.671 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 1383
Sequence: 1 MNNGKINGICALASGIALAG.....TDEVEAKKPKFDGVIXGW 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfill1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	775.5	56.1	289	US-09-489-039A-12337	Sequence 12337, A
2	757.5	54.8	279	US-09-543-681A-6534	Sequence 6534, Ap
3	685.5	49.6	285	US-09-489-039A-10641	Sequence 10641, A
4	671.5	48.6	279	US-09-543-681A-4548	Sequence 4548, Ap
5	651	47.1	272	US-09-489-039A-10410	Sequence 10410, A
6	596	43.1	279	US-09-543-681A-6914	Sequence 6914, Ap
7	520	37.6	118	US-09-540-236-2991	Sequence 2991, Ap
8	503.5	36.4	305	US-09-252-991A-18873	Sequence 18873, A
9	494.5	35.8	271	US-09-252-991A-20200	Sequence 20200, A
10	481	34.8	277	US-09-071-035-30	Sequence 30, App1
11	481	34.8	278	US-09-134-000C-4965	Sequence 4965, Ap
12	475	34.3	269	US-09-543-681A-4774	Sequence 4774, Ap
13	462	33.4	257	US-09-071-035-32	Sequence 32, App1
14	449.5	32.5	288	US-09-107-532A-7234	Sequence 7234, Ap
15	441	31.9	291	US-09-134-001C-4005	Sequence 4005, Ap
16	416	30.1	280	US-08-956-171F-5195	Sequence 5195, Ap
17	373.5	27.0	279	US-09-134-001C-4878	Sequence 4878, Ap
18	366	26.5	269	US-09-107-532A-6443	Sequence 6443, Ap
19	364	26.3	296	US-09-328-352-6924	Sequence 6924, Ap
20	364	26.3	326	US-09-489-039A-9512	Sequence 9512, Ap
21	358.5	25.9	261	US-09-634-238-420	Sequence 420, App
22	356	25.7	260	US-08-961-083-32	Sequence 32, App1
23	356	25.7	260	US-09-536-784-32	Sequence 32, App1
24	346	25.0	273	US-08-956-171F-5196	Sequence 5196, Ap
25	337.5	24.4	286	US-09-328-352-7028	Sequence 7028, Ap
26	335	24.2	304	US-09-328-352-7028	Sequence 7028, Ap
27	325.5	23.5	294	US-09-328-352-5549	Sequence 5549, Ap

28	324	23.4	272	US-09-071-035-22	Sequence 22, App1
29	322	23.3	342	US-09-252-991A-25420	Sequence 25420, A
30	321.5	23.2	280	US-09-489-039A-13968	Sequence 13968, A
31	319.5	23.1	220	US-09-634-238-283	Sequence 283, App
32	318	23.0	251	US-09-071-035-24	Sequence 24, App1
33	315.5	22.8	272	US-09-071-035-10	Sequence 10, App1
34	308.5	22.3	253	US-09-071-035-12	Sequence 12, App1
35	301.5	21.8	278	US-09-134-000C-5006	Sequence 5006, Ap
36	257.5	18.6	275	US-09-198-452A-291	Sequence 291, App
37	209	15.1	306	US-09-489-039A-10023	Sequence 10023, A
38	108	7.8	91	US-09-134-000C-3687	Sequence 3687, Ap
39	106	7.7	347	US-09-328-352-6764	Sequence 6764, Ap
40	105	7.6	794	US-09-134-000C-5516	Sequence 5516, Ap
41	103	7.4	503	US-09-071-035-360	Sequence 360, App
42	103	7.4	1074	US-09-071-035-358	Sequence 358, App
43	103	7.4	1074	US-09-071-035-394	Sequence 394, App
44	103	7.4	1096	US-09-134-000C-5764	Sequence 5764, Ap
45	102.5	7.4	308	US-09-134-001C-4400	Sequence 4400, Ap

ALIGNMENTS

RESULT 1
US-09-489-039A-12337
; Sequence 12337, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12337
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12337

Query Match	56.1%	Score 775.5	DB 4	Length 289
Best Local Similarity	55.2%	Pred. No. 3.3e-73		
Matches 153	Conservative 45	Mismatches 68	Indels 11	Gaps 3
QY	2	MNNGKINGICALASGIALAGCSNOSNEPAAISTTAAQTIKGVMAAGPEQAVAVACQVAK	61	
DB	22	NKTFPAVAGALIGSLALVGCQDEKDP-----NHIKGVIVGAHQVAVAAQVAK	73	
QY	62	KYNLIVELVEFNDYAMPNSAVSKGELDANAWOHKPYLEKSOCEKLANLIVYGTFTVPL	121	
DB	74	KYGLDELVELTFTNDYVLPENALSKGIDIVVAFQHKPYLLDQOIKDRGY-KLIVAGNTFTVYPI	132	
QY	122	AGSTTKITLNLKOGATTAVPNDPSNLARAILLEKQGLIKLKNNTNLFSTLDIVENP	181	
DB	133	AGSKKIKSLDELQPSQIAVPNDPTNLGRSLLLQOVGLIKLKGVLPLPSLIVENP	192	
QY	182	KKLVKEVTSVAARAIDD--VDLAVNNNVYAQVGLTASENGVEVEDKSPYNIIVAR	239	
DB	193	KULKIETELAPQPLPSLDQAIALAVINTTYSQGLTFPAKDGIFVEGKESPYNLIIVAR	252	
QY	240	ADNKSKAIQDPYKAYQDPEVAAEAKKPKFDGVIXGW	276	
DB	253	EDNKAENVKRFVQAVQSDVEVAAANKIFNGAVIXGW	289	

RESULT 2
US-09-543-681A-6534
; Sequence 6534, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:

APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 6534
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6534

Query Match 54.8%; Score 757.5; DB 4; Length 279;
Best Local Similarity 52.7%; Pred. No. 2.4e-71;
Matches 147; Conservative 49; Mismatches 70; Indels 13; Gaps 4;

QY 1 MFPERKINGICALASGIALAGCSNOSNEPAISKTAAQTIKYGVMAGPEQVAEYAGQVAK 60
DB 11 LKFKSLAVVSAIVGALALAGCGEKEKDP-----NHIRGVISGSEQOVAEYAKQVAK 62
QY 61 EKVMTLVLEFENDYAMNSAVSKGELDANAMQHKPYLEKDSQEKGLN-ILVIVGNTFVY 119
DB 63 DRYGLNVELVTFNDFVMPNESLSRGDIDINAFQHKPPL--DQQTIDKNYKITAAGNTFIY 120
QY 120 PLAGYSTKIKITLNEIKDQATIAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVE 179
DB 121 PLAGYSKKTITPLADLPDQAVAIIPNDPTNIGRSILILEKYGVLVLEKGVGLPTKDIIE 180
QY 180 NPKVLIVETVTSVAARAID--VDLAVNNNNYAGVGLTASENGVEVEKDSFYNIIV 237
DB 181 NPKVITVELAPQLPSPSLDQKLYLAVINTTASQVNLTPAKGKIFVEDKDSFYNIIV 240
QY 238 ARAADKSKAIQDFVKAQVQTEVEAEAKKQFKDGVIGW 276
DB 241 AREDNKODENKRTIQSYQTEDEVSAANKIFNGAVVGMW 279

RESULT 3
US-09-489-039A-10641
; Sequence 10641, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO: 10641
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10641

Query Match 49.6%; Score 685.5; DB 4; Length 285;
Best Local Similarity 50.6%; Pred. No. 9.4e-64;
Matches 135; Conservative 53; Mismatches 68; Indels 11; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAAQTIKYGVMAGPEQVAEYAGQVAKKYLTVELVE 71
DB 28 MLAGLALVGDQKGN-----AKHIVGVINGAEQDAEYAKKVAEKIGLEVLVG 79
QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNIVIVGNTFVYPLAGYSTKIKTL 131
DB 80 FSGSLLENSTAGDIDANVQHRPFLQDNKAHNV-HLVAAGNTVFPVPMAGYSRTIKSV 138
QY 132 NELKDGTIAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVEPKLVIVETV 191

DB 139 AELKDGTIAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVEPKLVIVETV 198
QY 192 SVARAID--VDLAVNNNNYAGVGLTASENGVEVEKDSFYNIIVARADNKSQAIQ 249
DB 199 AQLPRLDPPKVDVAIIISTYIQGTGLSPVRDGFIEDKNSPYNIIVTREDNKAENVK 258
QY 250 DFKAYQTEVEAEAKKQFKDGVIGW 276
DB 259 EFWQSTQSPPEVAKAAETIFNGAVVGMW 285

RESULT 4
US-09-543-681A-4548
; Sequence 4548, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 4548
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4548

Query Match 48.6%; Score 671.5; DB 4; Length 279;
Best Local Similarity 50.6%; Pred. No. 2.7e-62;
Matches 135; Conservative 49; Mismatches 72; Indels 11; Gaps 3;

QY 12 LASGIALAGCSNOSNEPAISKTAAQTIKYGVMAGPEQVAEYAGQVAKKYLTVELVE 71
DB 22 LLAGTLLTGCDNNDRYA-----IKVAVINGAEQDAEYAKKVAEKIGLEVLVG 73
QY 72 FNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKGLNIVIVGNTFVYPLAGYSTKIKTL 131
DB 74 FSGSLLENSTAGDIDANVQHRPFLQDNKAHNV-HLVAAGNTVFPVPMAGYSRTIKHP 132
QY 132 NELKDGTIAVPNDPSNLARALLLEKQGLIKLKDNTNLFSTLTDIVEPKLVIVETV 191
DB 133 QELKGGDITIALPNDPTNIGRALLLLEKQGLITLNPDSGLPTLTDIDNPDKINIMOLEG 192
QY 192 SVARAID--VDLAVNNNNYAGVGLTASENGVEVEKDSFYNIIVARADNKSQAIQ 249
DB 193 AQLPRLVNDPPVTAIIISTYIQGTGLSPVRDGFIEDKNSPYNIIVTREDNKAENVK 252
QY 250 DFKAYQTEVEAEAKKQFKDGVIGW 276
DB 253 DFKAYQSPPEVATAETIFKGAIVGMW 279

RESULT 5
US-09-489-039A-10410
; Sequence 10410, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO: 10410
; LENGTH: 272

Page 3

	Query March	43.1%	Score 596;	DB 4;	Length 279;
	Best Local Similarity	42.1%;	Pred. No.2.4e-54;		
	Matches 112;	Conservative	62;	Mismatches 78;	Indels 14; Gaps 4;
QY	14 SGIALAGCS-NQSNEPPAISKTAQTIKVGYMAGEBQAVAEVAQVANAKXNLTVELEEF	72			
	: : : : : :	:	:	:	:
Db	25 TSLVLATGCSPEEGDKP-----LKVALNTGPDOQIMDEBVYKLAKEOGSLDIKVITF	74			
QY	73 NDVMPASAVSAGKGLDANAMGHKPYLEKDSEOKGNINLVIGNTGVYFLAASYRKITLN	132			
	: : : : : :	:	:	:	:
Db	75 NDYLPLEALRNNDVDVNAAFOTVPYYEKSGERGY-QETLIGKTFFPIAYNSNKITIE	133			
QY	133 ELKDGATIAVENDSNTARALLILEKOGILIKLKNNTNLFSTLLDIVENPKLVIKEVDTS	192			
	: : : : : :	:	:	:	:
Db	134 ALPGGATVAIISNEGATTGRSILLTQAQGLIKLKGGVGLPTTLIDIENTPKLKFEADITP	193			
QY	193 VAAAIAD--VDLIAVVNNVAGOVGLTASENGVEVEDSDRYNNTIVRARMKSKAIOD	250			
	: : : : : :	:	:	:	:
Db	194 QLTLLSDPNLYLSIIINNFSOVLAKRKGLEFWENTDSFYVNLIVARAIDKONERLUK	253			
QY	251 FVKAYQTDVEAEAKKOFKDGVIKGM	276			
	: : : : : :	:	:	:	:
Db	254 LVAVFOSEDILOKAOEVYKGDVAVKAM	279			

Query Match	36.4%;	Score 503.5;	DB 4;	Length 306;
Best Local Similarity	43.0%;	Pred. 1.56-44;		
Matches 107;	Conservative 51;	Mismatches 84;	Indels 7;	Gaps 3;

QY	29	AAISKTAQGITKYGVAAGPEQAAVAEAVAGQV---AKKKNLTVELVEPNDYAMNSAVSK	84
Db	58	AAAGLTAAQAESITVAATPEVPAHEIINVVKPLTAKGVDLTKT-EFTDVPQGNVQVSE	115
QY	85	GELDANAMQHKPYLEKDSOEKLNINLVITVGNTEFVYPPLAGYSTKIKTINELKDGGATIAVFN	144
Db	116	KRLDANFQHQPLIDENFKAG--TDLVAIVGVHIEPLGAASSKKKKLDDELPSGATIVIPN	174
QY	145	DPSNLAARLILILEKQGLIKLKDQNTNLFSTLLDIYENPKYIKVETDSVVAARAIDDDVLA	204
Db	175	DATNGGAILLLLPKAGVIKIKDKKSLTATPRDVLVDNKKTKIRLELAATLPRVLTQVDMA	234
QY	205	VANNNVAGQVGLPASENGVFEVEDKDSPEYVNIIVAAPDNKSKAIQDPEVKVAVQTDVEAEA	264

```
78;
12; Gaps 5
```

QY 3 FKGINGICALAGSAGSNOSEPAISKTAAQTIKVGWAGPEQVAEVAQVAKK 62
DB 5 FSKIGLIGVLA-FTIACASGS-----VKDTKETVKTGVGKNDWESVMDRLKK- 57
QY 63 YNLVELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSOEKLNULVYVGTFTVPLA 122
DB 58 -NIDQLVEFTDYQPNALAEKEBIDINAFQHOIFLDNNKHEG-TKLVSIGNTVAPLG 115
QY 123 GYSTKIKTINELKOGATIAVNDPSNLABALILEKOGILTKDNTLFTSLDIVENPK 182
DB 116 IYANKDKITIKOGELIAPNDPTNGRRLILQTAGLIVDPAKQOLPVSDITENKR 175
QY 183 KLVIKEVDTSVAARIDVDLAVNNNVAQVGLTASENGVE--DKDSPYVNIIVAR 239
DB 176 QKITEIDATQARALQVDASVINGVADAGTDPKDALFLPVENKAKPYVNIIVAR 235
QY 240 ADMNDKAIQDPVAKYQVDEVEAEAKKQFKDGVIKGM 276
DB 236 EEDQENKLYQVVEYQOEETKVIATSKGANVPAM 272

RESULT 12
US-09-543-681A-4774
Sequence 4774, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 4774
LENGTH: 269
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-4774

Query Match 34.3%; Score 475; DB 4; Length 269;
Best Local Similarity 35.8%; Pred. No. 1,2e+41;
Matches 96; Conservative 60; Mismatches 104; Indels 8; Gaps 4;
QY 12 LASGIALAGCSNOSEPAISKTAAQTIKVGWAGPEQVAEVAQVAKKINTLVE 71
DB 7 IITGLIASISLFS---VSCSQEDIDYVKAINTGPDQVIMDEVIRLAKTEGLNVEVIA 63
QY 72 FNDVAMPNSAVSKGELDANAMQHKPYLEKDSOEKLN-NUYVONTFTVPLAGSTIKT 130
DB 64 FNDVERPKALENRVDVNFQSTIPYD--SEIRAHNYKFKHTASKTIHFFPLAYSKISD 121
QY 131 INELKDGATIAVNDPSNLABALILEKOGILTKDNTLFTSLDIVENPKLVIK 190
DB 122 ISELVEGVVAIPNDSKGRGALLILAEHNLISLKEGVGLPSEVEDIIDNNALIFHEVE 181
QY 191 TSVARALDD--VDLAVNNNVAQVGLTASENGVEVDKSPYVNIIVARADKDSKAI 248
DB 182 TPLMVEALDDPEVTMAVINNNFSQIGLATRDGLIMENKSPYANVVVTRIDNNDEKI 241
QY 249 QDFVAKYQVDEVEAEAKKQFKDGVIKGM 276
DB 242 KKLITVJHSRQVELKVGEMTKGDVAKAM 269

RESULT 13
US-09-071-035-32
Sequence 32, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: G11 H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-32

Query Match 33.4%; Score 462; DB 4; Length 257;
Best Local Similarity 38.0%; Pred. No. 2.6e+40;
Matches 95; Conservative 56; Mismatches 93; Indels 6; Gaps 3;
QY 30 ASKTAQTIKVGWAGPEQVAEVAQVAKKINTLVEFNDYAMPNSAVSKGELDA 89
DB 5 SVMDTEITVKTGVGVTNDWESVMDRLKK--NIDQLVEFTDYQPNALAEKEIDL 62
QY 90 NAMQHKPYLEKDSOEKLNULVYVGTFTVPLAGYSTIKTINELKDGATIAVNDPSNL 149
DB 63 NAFQHOIFLDNNKHEG-TKLVSIGNTVAPLGIVANKDKITIKDGEIAPNDPTNG 121
QY 150 ARALILEKOGILTKDNTLFTSLDIVENPKLVIKEVDTSVARAIDVDLAVNNN 209
DB 122 GRLLILQTAGLIVDPAKQOLPVSDITENKQKLTBLDNTQARALQVDASVING 181
QY 210 YAGQVGLTASENGVE--DKDSPYVNIIVARADKDSKAIQDFVAKYQVDEVEAEAKK 266
DB 182 MAVDAGYTPDKALIFLEPVNERAKPYVNIIVAREEOENKLYQVVEYQOEBETKVI 241
QY 267 QFKDGVIKGM 276
DB 242 TSKGANVPAM 251

RESULT 14
US-09-107-532A-7234
Sequence 7234, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street

CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107, 532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 7234:
SEQUENCE CHARACTERISTICS:
LENGTH: 291 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...291
SEQUENCE DESCRIPTION: SEQ ID NO: 7234:
US-09-107-532A-7234

Query Match 32.5%; Score 449.5; DB 4; Length 291;
Best Local Similarity 37.9%; Pred. No. 6.6e-39;
Matches 100; Conservative 51; Mismatches 104; Indels 9; Gaps 4;
QY 16 IALAGCSNQSNEPAISKTAAQTIKVGWAGPEQAVAEVAGVAKKXNLTVLVEFNDY 75
Db 28 IWLACSGKADTDKKT--IVKLGIGEDTDVMDVKDLKKE--NINLKVYKFTDY 82
QY 76 AMPNSAVSKGELDANAMOHKPYLEKDSQKGLNLTIVGNTFVYPLAGYSTKIKTLNEK 135
Db 83 NOPNVALDGSIDLSFQHPFLDNYNKEHG-TDLVSGTVAAPLGIYSTLTLDVADIK 141
QY 136 DGATVAVPDSNLARALILLEKQGLIKLKDNTNLFSTLTDIVENPKKLVKEVDTSVAA 195
Db 142 DGAKVAIPNDVNGRALLLLQTAGLIVDKAKGQAPVSDITENRKNLDSILDASQTA 201
QY 196 PAIDVDVLA VNNNAGVGLTASENGVFEVDKDS---PYNITVAPADNKSFAIDFV 252
Db 202 RALSDVDIVSINSGAVAGFNPASDALEFPVDHARFYNIIVAKKEENNTKYOKIV 261
QY 253 KAYOTDEVEAEKKQKQFQKGVTKGW 276
Db 262 DAYOTEDTVKVIKTSKSSVPAW 285
RESULT 15
US-09-134-001C-4005
Sequence 4005, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134, 001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064, 964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055, 779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4005
TYPE: PRT
LENGTH: 288
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4005
Query Match 31.9%; Score 441; DB 4; Length 288;
Best Local Similarity 36.2%; Pred. No. 5.1e-38;
Matches 94; Conservative 55; Mismatches 95; Indels 16; Gaps 4;
QY 16 IALAGCSNQSNEPAISKTAAQTIKVGWAGPEQAVAEVAGVAKKXNLTVLVEFNDY 75
Db 20 IWLASAGGNN-----GKVTIGVANSNTKAMKVK-ELAK-KODIDLEIHFSFY 67
QY 76 AMPNSAVSKGELDANAMOHKPYLEKDSQKGLNLTIVGNTFVYPLAGYSTKIKTLNEK 135
Db 68 NPENKALSDGDDIDLNFQHFAPLDQYKKAHKDNTIEALSTVLAFLGIYSDKVKNIKQVK 127
QY 136 DGATVAVPDSNLARALILLEKQGLIKLKDNTNLFSTLTDIVENPKKLVKEVDTSVAA 195
Db 128 KGAQVAIPNDVSNQARALILLESAGLIKQKQFGLNGTTKDIENPKDKLTVAVDAQOTA 187
QY 196 PAIDVDVLA VNNNAGVGLTASENGVFE---DKSPYNIIVAPADNKSFAIDFV 251
Db 188 RALSDVDIVSINNGVATKAGDAKDPYLEKASSDAVAPYINVAVANSKDKDKNTYKXI 247
QY 252 KAYOTDEVEAEKKQKQFQKGVTKGW 271
Db 248 IELVHSKKAQKALKEEDTKDG 267

RESULT 16
US-08-956-171E-5195
Sequence 5195, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gail H. Choi
Patrick S. Dillon
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956, 171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009, 861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/761, 986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5195:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5195:
US-08-956-171E-5195

Query Match 30.1%; Score 416; DB 4; Length 280;
Best Local Similarity 35.8%; Pred. No. 2,1e-35;
Matches 93; Conservative 50; Mismatches 101; Indels 16; Gaps 4;

QY 16 IALAGCNQNEPRAISKTAQITKVGWAGPEQAVAEVAGVAKKYNLTVELVENDY 75
DB 13 VLLAACGNDND-----KKVTIGVASNDTGAKEVK-ELAK-KDDIDVEIKHSY 60
QY 76 AMPNSAVSKGELDANAMQHKPYLEKDSQEKLNLTIVGNTFYVPLAGYSTIKITNELK 135
DB 61 NLPRKALNDGIDWNAFQHFAPLDQYKKAHKGTISALSTTVLAPLGYSDKIKDYKVK 120
QY 136 DGATIAVNPDSNLRALLILEKQGLIKLKDNTNLSTTLDIYENPKVYIKYDVSVAA 195
DB 121 DGAKVVIPIPVSNQARALKLEAAGLIKTKDFELAGTVKDITSNPHLKITVDAQOTA 180
QY 196 RAIDVDLAVVNNYVAGVGLTASENGVEYEDKDS---PYVNIIVARADNKSKAIQDF 251
DB 181 RALSDVIAIYVNSGVATKAGCDPRNDPIFEKNSDAVKRYINIVAVNDLDNKTAKI 240
QY 252 VKAYQTDVEAEAKKQFQDG 271
DB 241 VELYSKKAQKALQEDVDG 260

RESULT 17
US-09-134-001C-4878
Sequence 4878, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4878
LENGTH: 279
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4878

Query Match 27.0%; Score 373.5; DB 4; Length 279;
Best Local Similarity 33.9%; Pred. No. 6,1e-31;
Matches 92; Conservative 47; Mismatches 115; Indels 17; Gaps 7;

QY 9 ICALASGIALAGCNQNEPRAISKTAQITKVGWAGPEQAVAEVAGVAKKYNLTVELV 67
DB 17 ILVALITLALACG-----GKDKKITTVGASAPFAHEILBKAKPLIKKGYDLKI 68
QY 68 ELVEFNDYAMPNSAVSKGELDANAMQHKPYLEKDSQEKLNLTIVGNTFYVPLAGYSTK 127

DB 69 KPI--NDVTPNKLLDKGEIDANFPQHTPIYNTESKEGY-KISAGNVELEPMVAESK 125
QY 128 ITTLNELKDGTIANVPNDSNLRALLILEKQGLIKLKDNTNLSTTLDIYENPKVI 186
DB 126 YSLDLPKGAIVVYVSNPABEQRFLEKFEVDEGLIKLKGVKIEAKFDDITENKDIKF 185
QY 187 KEVDTSVABAI---DVPDLAVVNNYVAGVGLTASENGVEYED-KDSPYNIIVARADN 242
DB 186 NKGSAEYIKRYQNDADAVIYNSNVAIDQLSPKXDSINLESKDPYANLIVKKGH 245
QY 243 KDSKAIQDFVKAAYQTDVEAEAKKQFQDGYI 273
DB 246 KDDKNIKVLMETVLQSKETQIDYIKDYGDVAV 276

RESULT 18
US-09-107-532A-6443
Sequence 6443, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham

STATE: Massachusetts
COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6443:

SEQUENCE CHARACTERISTICS:

LENGTH: 269 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

HYPOTHETICAL: YES
ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium
FEATURE:

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...269

SEQUENCE DESCRIPTION: SEQ ID NO: 6443:
US-09-107-532A-6443

Query Match 26.5%; Score 366; DB 4; Length 269;
Best Local Similarity 35.5%; Pred. No. 3,5e-30;
Matches 89; Conservative 43; Mismatches 85; Indels 34; Gaps 7;

QY 42 GYAGPEQAV-----AEVAGV-----AKKYNLTVELVENDY 74
DB 3 GRNGPQQLNRIPIVSCETTPPGALCSAPISPAEMGDHVNPLWTHGFDL--ETGFGDD 60

```

QY 75 YAMNSAVSKGELDNANQHKRYLEKDSQEKLNVLVNTFVYPLAGSTKIKTINEL 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 YLPNGSLQGGIDIDANYPQHIPLYNKEIKERY -DFVAGVNHLEPMGLSKKVDSEL 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 KOGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTTL-DIVENPKLVKE-VDTG 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 KOGATITNSSESDBGRITITLQDAGLVKVDGVDLEFATDDLEENPKDKFHHITDPS 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 VAAARAI--DDVDLAVNNNYAGQVGLTASENGVEEDKOSPYVNIIVARADNKSXAIO 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 LLAFTYQNDGDLVAINANFAYGAGLNPLEDAVLLEKDNSPYVNIIVARADNKSXAIO 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 251 FVKAYOTDEVE 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 LIEVAKREVO 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 19
US-09-328-352-6924
; Sequence 6924, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6924
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6924

```

Query Match 26.3%; Score 364; DB 4; Length 296;

Best Local Similarity 36.3%; Pred. No. 6.7e-30;

Matches 99; Conservative 45; Mismatches 115; Indels 14; Gaps 7;

```

QY 8 GICALASGIALAGCNSQNSNEPAISKTAQITIKVGMAGPEQAAVEAGVAAEKYLV 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 33 GITLIVLGLVGRYIKN-----KQADVLITIGISPPYAEHLQSVANEV--EKQGVH 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 ELVEFNDYAMNSAVSKGELDNANQHKRYLEKDSQEKLN-NLIVGNTFVYPLAGYST 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 KLVEFSDVRAPNIAVONGDIDANFQOSVFLRNAVKETGYDLHAFGVSGSHVGL--YSK 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 KIKTINELKOGATIAVPNDPSNLARALILEKQGLIKLKDNTNLFSTTLIVENPKLV 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 143 KYKSLIEDPLNARVAIRPDVNLARALILHRAGLIQKQINNELSTTQIILANPKQLS 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 KEVDTSVARAIDVDVLA VNNNYAGQVGLTASENGVEEDKOSPYVNIIVARADNKS 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 203 VEVEGPGCAHAAYNDVDLIFGPHYLKMAKVTDPHSALFLDPIDKVAILEFVTRRDYQDKN 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 246 KAIODPVKAYO-TDEVEAEAKKQPKDGV-IKGW 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 QKLAIFVAFQNSKQADQIILDKDFGKGMFQGW 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 20
US-09-489-039A-9512
; Sequence 9512, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747

```

```

; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9512
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9512

```

Query Match 26.3%; Score 364; DB 4; Length 326;

Best Local Similarity 35.3%; Pred. No. 7.8e-30;

Matches 94; Conservative 49; Mismatches 105; Indels 18; Gaps 6;

```

QY 14 SGIALAGCNSQNSNEPAISKTAQITIKVGMAGPEQAAVEAGVAAEKYLV 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 AGVALS-----LSLWMTSAQAAALRVADPVPH-ABILNLYIKKIDPSLDLVVELT 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 DYAMNSAVSKGELDNANQHKRYLEKDSQEKLNVLVNTFVYPLAGSTKIKTIN 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 SGVANNELLASGDVANYFQHVPLYK--DQEKALGKFTVAATVAIIEPLGIYSHKXDFS 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 ELKOGATIAVPNDPSNLARALILEKQGLIKL---KDNTNLFSTTLIVENPKLV 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 SLPERATVAVPNTNLSKALFLQAKLIKDPFTTPATTLATPKDIVENPKLIKLE 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 189 VDTSVARAIDVDVLA VNNNYAGQVGLTASENGVEEDKOSPYVNIIVARADNKS 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 IESPOIPSLDDVDVLA VNNNYAGQVGLTASENGVEEDKOSPYVNIIVARADNKS 294
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 248 IODPVKAYOTDEVEAEAKKQPKDGV 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 IKALAKDLTSPQVAEFIRKTINGSVI 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 21

US-09-634-238-420

; Sequence 420, Application US/09634238

; Patent No. 6544772

; GENERAL INFORMATION:

; APPLICANT: Glenn, Matthew

; APPLICANT: Havakala, Ilkka J.

; APPLICANT: Bloksberg, Leonard, N.

; APPLICANT: Lubbers, Mark W.

; APPLICANT: Dekker, James

; APPLICANT: Christensen, Anna C.

; APPLICANT: Holland, Ross

; APPLICANT: O'Toole, Paul W.

; APPLICANT: Reid, Julian R.

; APPLICANT: Coolbear, Timothy

; TITLE OF INVENTION: Polynucleotides, materials incorporating

; FILE REFERENCE: 11000.1043U

; CURRENT APPLICATION NUMBER: US/09/634.238

; CURRENT FILING DATE: 2000-08-08

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 420

; LENGTH: 261

; TYPE: PRT

; ORGANISM: Lactobacillus rhamnosus

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(261)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-634-238-420

Query Match 25.9%; Score 358.5; DB 4; Length 261;

Best Local Similarity 36.3%; Pred. No. 2.1e-29;

Matches 91; Conservative 52; Mismatches 91; Indels 17; Gaps 9;

```

QY 12 LASGIALAGCNSQNSNEPAISKTAQITIKVGMAGPEQAAVEAGVAAEKYLV 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 LSSWRRAKSSSSS-----SKN--ETVKGIMSTDKIWTIDIOSRLKQ--GVTIKIVQ 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


Db 187 NFEVTEAKLDYKKS-LFEKQADENSQWYNIIVAKDMETSPKDAIKVIAAHTTDDVK 245
QY 262 AEAKKQFKGV 272
Db 246 -KVIESSDGL 255

RESULT 24
US-08-956-171E-5196
; Sequence 5196, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesh
; Gail H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: P8348P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 5196:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5196:
US-08-956-171E-5196

Query Match 25.0%; Score 346; DB 4; Length 273;
Best Local Similarity 31.0%; Pred. No. 4.6e-28;
Matches 85; Conservative 53; Mismatches 116; Indels 20; Gaps 8;

QY 9 ICALASGIALAGCSNQNNEPAISKTAQTIKYGVWAGPEQVAEVAAGVAKER-YNLTV 67
Db 8 ILVLTFAVVLAAACNGK-----SGSDDKITVASAPPAHEILEKAKPILLEKGYELDI 62

QY 68 ELVEFNDYAMPNSAVSGELDANAMOKHFKYLEXDSQSKGLNNIVYIGNTVYVLAGSTK 127
Db 63 KTI--NDYTPNKLIDGSEIDANTFQHTPIHTEKDKGY-KIVSAGDVHLEPMAYVSKK 119

QY 128 IKTLNEKDGATIAVPDPSNMLARALILEKQGLIKLKDNTN-----FSTTLDIVENPKK 183

Db 120 YKSLKELPKGATVYVSNNEPAEQGFLEFVDAGLIKIKKGVKIEDAKFS---DITENKDD 176
QY 184 LVIEKVDTSVAARAI---DDVDLAVVNNYVAGGVGLTAENGVFVED-XDSFVYNIIVAR 239
Db 177 IKFNKQSAEFLPKIYQNEBDADAVIINSFPAIEQKLNPKDSIIVASADNPYALILAVK 236

QY 240 ADNKSRAIODFVAVYQTDVEAEAKQFKGYI 273
Db 237 EGHDDKKIKALIEVLQSKDIODFINEKYNGAVI 270

RESULT 25
US-09-328-352-6702
; Sequence 6702, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6702
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii

US-09-328-352-6702

Query Match 24.4%; Score 337.5; DB 4; Length 286;
Best Local Similarity 37.7%; Pred. No. 3.9e-27;
Matches 87; Conservative 40; Mismatches 89; Indels 15; Gaps 7;

QY 33 KTAAGTIKGV---MAGPEQVAEVAAGVAKERKYLTEIVERNYAMPNSVSGELDA 89
Db 41 QGSENELVGISPPPAKPLQAAADA-----RKQGLNKLVEFSQWNTPTTLNGLDIDA 95

QY 90 NAMQHKYLEXDSQSKGLN-NLIVIGNTFVYPLAGYSTIKITLNEKDGATIAVPDPSN 148
Db 96 NFFQHQPLDNKAIKGTGFLKAFVAGASHGL--YSKKYSLDLPQNAVAVIPNDVYN 153

QY 149 LRALILIEKQGLIKLKDNTNLFSTTLDIVENPKLVIKVDTSVAARAIDDVDLAVNN 208
Db 154 QSRALLLQQAFLITLKDSNNHLSALKDVSNPKRLCFIEVEGPOTARAIIDVDLAFGYP 213

QY 209 NVAGQVGLFASENGVFVEDKDS--PYNNIIVARANKR-SVAIDOFVAYQ 256
Db 214 HTL-RLAKTADPNSALLDDNTNKRVAIFVVRDYEKDKLKKFVEIYQ 263

RESULT 26
US-09-328-352-7028
; Sequence 7028, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7028
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii

US-09-328-352-7028

Query Match 24.2%; Score 335; DB 4; Length 304;
Best Local Similarity 32.0%; Pred. No. 7.9e-27;
Matches 88; Conservative 43; Mismatches 126; Indels 18; Gaps 5;

QY 5 KINGICALASGIALAGCSNQNNEPAISKTA-AQTIKYGWAGPEQVAEVAAGVAKERY 63

Tue Jun 22 11:04:04 2004

us-10-018-672-2.rat

Page 11

Db 19 KTLISFLSVSVLLAACKQONFQNGKDSAOLOTVAISTGSPADJWRIRIATLPTKAA 78
Qy 64 NLTVLVEFENDYAMPNSAVSKGELDANAMQHKPYL-----EKDSQEKLNNTVIYGNTP 117
Db 79 GLKLEVKNFETDYAMNTATANKVEDLNAFCYAVLVAFNQSKD-----KIAPVATY 131
Qy 118 VYPLAGYSTKIKITNELKDGTAVPNDPSNLARALLLEKQGLIKL-KONTLFTTLD 176
Db 132 LEPWGIYSKTKKVDPEPQASIAIPNDANEARALLLSQAGLIKAKADFPDPAKGPSPD 191
Qy 177 IVENPKLVIKEVDTSVAARAIDVDVLA VNNNNAAGVGLTASENGVEFEDKSP---YV 233
Db 192 ITDSSKKIDIKPIGMATATVRKDEVDALVIGNTLAMGGLNVLKDSIYEPVQSTKLV 251
Qy 234 NIIVARADNKDSKAIQDFVKAQYQDEVEAEAKKQF 268
Db 252 NILATASRKDDPVLQKVGQLYHTEAVKXYEQHF 286

RESULT 27
US-09-328-352-5549
; Sequence 5549, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Brelton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5549
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5549

Query Match 23.5%; Score 325.5; DB 4; Length 294;
Best Local Similarity 32.9%; Pred. No. 7.4e-26;
Matches 91; Conservative 48; Mismatches 119; Indels 19; Gaps 8;
Qy 5 KINGICLAGGIALAGCSNOSPRAISKTPA-----OTIKVWAGPEQAVAE-VAG 56
Db 6 KLFVFLRSASVLTTLTACNKQ---PAQTEMTAAKDKTESVRTIKL-VSTGSDTDVWKYVAT 61
Qy 57 QVAREKYNLTVELVEFNDYAMPNSAVSKGELDANAMQHKPYLER-DSQEKGLNNLVYGN 115
Db 62 LPETRAQGITKLEVTNLDYVLTNTSVASGEODVNAFQSFYTLAAYNASNA--KVAAVAT 119
Qy 116 TVVYPLAGYSTKIKITNELKDGTAVPNDPSNLARALLLEKQGLIKL-KONTN-LFSTT 174
Db 120 TYLEPMGIYANKVTVDEFPQAGATIAIPNDTANEARALLTLOSQGLIKL-KPFPDPAKGP 179
Qy 175 LDIYENPKLVIKEVDTSVAARAIDVDVLA VNNNNAAGVGLTASENGVEFEDKSP--- 231
Db 180 NDIENENPKAIQKPIQMTTAVRVKDDVDAVLGNTLMBGLNVMKDAIFRSPIDSTKL 239
Qy 232 YNVIIVARADNKDSKAIQDFVKAQYQDEVEAEAKKQF 268
Db 240 YVNLIGVAEANKDDPIYTKLGEVLHLPKQKFNVEKF 276

RESULT 28
US-09-071-035-22
; Sequence 22, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 272 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-22

Query Match 23.4%; Score 324; DB 4; Length 272;
Best Local Similarity 30.3%; Pred. No. 9.4e-26;
Matches 81; Conservative 56; Mismatches 104; Indels 26; Gaps 8;
Qy 16 IALAGCSNOSPRAISKTPAQTIRKVGWAGPEQAVAEVAGQVAK-EKYNLTVELVEFN 73
Db 17 LVIGAGCSNKKSDS-----VLKVGASPVPH---ABILEHVKKLLEKGVKLEVTYTT 65
Qy 74 DYAMPNSAVSKGELDANAMQHKPYLERKDSQEKGLNNLVYGNTPFYPLAGYSTKIKITNE 133
Db 66 DYVLPKKALESIGDIDANFYOHVPFENEAVKENDY-DFVNAAGIHLEPGLYSKKYKSIOE 124
Qy 134 LKDGATIAVPNDPSNLARALLLEKQGLIKL-KONTNLFSTTL-DIVENPKLVIK-EVDT 191
Db 125 IPDGSSTIYSSSVSDMPRVTLTLEDAGLITLKEGVDRTRTATDDIDKTKLKRFBEDP 184
Qy 192 SVAARAIDVDVLA--VNNNNAAGVGLTASENGVEFEDKSPYNIIVARADNKDSKAIQ 249
Db 185 AIMTTLVDEEAGAAVLINSNFAVDGLNPKKDAIALEKSSPYANIIIVAREDENNERVK 244
Qy 250 DFVKAYQDEVEAEAKKQFKQGVIKGW 276
Db 245 KLVKVLRSKEVO-----DWITKKW 263

RESULT 29
US-09-252-991A-25420
; Sequence 25420, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-24

Query Match 23.0%; Score 318; DB 4; Length 251;
Best Local Similarity 30.5%; Pred. No. 3,5e-25;
Matches 80; Conservative 55; Mismatches 101; Indels 26; Gaps 8;

QY 21 CSNQSNEPAISKTAQTIKYGVNAGPEQAVAEVAGQVAK--EKXNLTVELVEFNDYAMP 78
DB 1 CGNKSDSDS-----VLKVGASPVPR---ABILIEHKPLLEKGVLTETTTDYLP 49
QY 79 NSAYSKGLDANAMQHKRYLEKDSQEKGLNLVYGNFTVYPLAGYSTKIKTLNELDGA 138
DB 50 NKALSGGIDANYPQHVPFFNEAVKENDY-DFVNAAGAIHEPVLGYSKYSLQEIPIDGS 108
QY 139 TIAVNDPSNLARALLILEKQGLIKLKDNTLFTTL-DIVENPKKLVIK-EVDTVAAR 196
DB 109 TIYSSSVSDMPRLVLTIEDAGLTLKGVRTATPDIDDKNTKKLKNHESDPALMTT 168
QY 197 AIDVDLA--VNNNYAGQVGLTASENGVEFVEDKSPYNNIIVARADNKSKAIQDFVKA 254
DB 169 LYDNEEGAAVLINSNFAVDQGLNPKDAIALEKSSPYANNIIVARKDENNENKLVKY 228
QY 255 YQDVEVEAEAKQKPFQGVIKGM 276
DB 229 LRSKEVQ-----DWITTKM 242

RESULT 33

US-09-071-035-10
Sequence 10, Application US/09071035
Patent No. 6448043

GENERAL INFORMATION:

APPLICANT: Gil H. Choi

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/071,035

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes

REGISTRATION NUMBER: 36,373

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 272 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-10

Query Match 22.8%; Score 315.5; DB 4; Length 272;
Best Local Similarity 31.9%; Pred. No. 7,4e-25;
Matches 84; Conservative 52; Mismatches 116; Indels 11; Gaps 7;

QY 11 ALASGIALAGCSNQSNEPAISKTAQTIKYGVNAGPEQAVAEVAGQVAKKXNLTVELV 70
DB 10 AVIATVILAACG--GNKQA--DQKEDKEITVAVOLSSKDIETIAKKA-EKKGYKINIM 64
QY 71 EFNQVAMPNSAVSKGLDANAMQHKRYLEKDSQEKGLNLVYGNFTVYPLAGYSTKIKT 130
DB 65 EYSDNVAINDAVQHDADANFQHPFEMFNKKA-DLVAQPIYYAGGFSKSEYD 123
QY 131 LNELKQATIAVNDPSNLARALLILEKQGLIKLKDNTLFTTL-DIVENPKKLVIK 190
DB 124 AKDLPENAKVGRPSDPTNEGRLALILNANGVILKKEGVFGVADVDPENPNITFESID 183
QY 191 TSVARAIIDVDLAVI--NNNYAGQVGLTASENGVEFVEDKSP--YNNIIVARADNKSK 246
DB 184 LNLAAVDEKDIAMFCYPALTEPAGLT-TKDAIILEKSKSKYVALQVTRKGEKSE 242
QY 247 AIQDFKAYQDVEVEAEAKQK 269
DB 243 KIKVLKEMTTKEVAEYIKNSK 265

RESULT 34

US-09-071-035-12
Sequence 12, Application US/09071035
Patent No. 6448043

GENERAL INFORMATION:

APPLICANT: Gil H. Choi

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/071,035

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-12

Query Match 22.3%; Score 308.5; DB 4; Length 253;
Best Local Similarity 33.2%; Pred. No. 3.6e-24;
Matches 78; Conservative 46; Mismatches 104; Indels 7; Gaps 5;

39 IKVGVAPGEOVAEVAAGVAKKXNLVLEFENDYAMPNSAVSKGELDANAMQHPYL 98
15 ITVAVQLESSEKDIIEIAKKA-EKKGYKINIMEVSDNVAVDVQCHDEADANFAQHQP 73
99 EKDSQKGLNNLVVGNFTVYPLAGYSTIKITLNEKOGATIAVNDPSNLARLLILEK 158
74 EWFNKKKA-DLVAQPIYFAGGFYSKEDADLPENAKVGIIPSDPTNKGRLATILNA 132
159 QGLITKXNTNLFSTLIVENPKLVKEVDTSVAARIDVDLAVV--NNYAGQVGL 216
133 NVVILKKGVGNGTVADVENPKITFESIDLNAQVDEKQIAVFCYPALEPAGL 192
217 TASENGVEVEDKSP--YVNIIVARADNKSRAIQDFVAYQTBVEAKKQK 269
199 T-TXQALLLEDKASKHYALGVVTRKGEKSEKIKVLEAMTKEVAEYIKKNS 246

RESULT 35
US-09-134-000C-5006
Sequence 5006, Application US/09134000C
Patent No. 6617156

GENERAL INFORMATION:
APPLICANT: Iyrm Doucet-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US/09/134,000C
PRIORITY FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patent version 3.1
SEQ ID NO 5006
LENGTH: 278
TYPE: PR1
ORGANISM: Enterococcus faecalis
US-09-134-000C-5006

Query Match 21.8%; Score 301.5; DB 4; Length 278;
Best Local Similarity 31.5%; Pred. No. 2.3e-23;
Matches 78; Conservative 52; Mismatches 107; Indels 11; Gaps 7;

11 AASGALAGCSNQSNEPAISKTAQTIKVGWAGPEQVAEVAAGVAKKXNLVTLV 70
14 AVIAVITLACG--GKKA--DQKEDKETTAVQLESSEKDIIEIAKKA-EKKGYKINIM 68
71 EENDYAMPNSAVSKGELDANAMQHPYLEKDSQKGLNNLVVGNFTVYPLAGYSTIKIT 130
69 EVSDNVAVDVQCHDEADANFAQHQPFWEMFNKKKA-DLVAQPIYFAGGFYSKED 127
131 LNEIKGATIAVNDPSNLARLLILEKOGITKXNTNLFSTLIVENPKLVKEVD 190
128 AKDLPENAKVGIIPSDPTNKGRLATILNANGVILKKGVGNGTVADVENPKITFESID 187
191 TSAARIDVDLAVV--NNYAGQVGLTASENGVEVEDKSP--YVNIIVARADNKS 246

DB 188 LNLAKVDEKDIAMVFCYPALEPAGLT-TKDALLLEDKASKHYALQVTRKGEKDS 246
QY 247 AIQDFVKA 254
DB 247 KIDGFRS 254

RESULT 36
US-09-198-452A-291
Sequence 291, Application US/09198452A
Patent No. 6559294

GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 291
LENGTH: 275
TYPE: PR1
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-291

Query Match 18.6%; Score 257.5; DB 4; Length 275;
Best Local Similarity 32.2%; Pred. No. 9.5e-19;
Matches 77; Conservative 49; Mismatches 98; Indels 15; Gaps 9;

33 KTAQGTIKVWAGPEQVAEVAAGVAKKXNLVLEFENDYAMPNSAVSKGELDAN 90
23 KEDAQ-KIRIVASPTP-AELLESQEAADGKIKLIPVDDIRIRRLLDQVDAN 80
91 AMQKPYLEKDSOE--KGLNNLVVGNFTVYPLAGYSTIKITLNEKOGATIAVND 145
81 YFQHQAFLEDECEERYDCKG--ELVVIKAVHLEPQATYSKHSLSRLKSOQKLTITAPVD 138
146 PSNLARLLILEKOGITKXNTNLFSTLIV--ENPKLVKEVDTSVAARIDVDL 203
139 RTVAQALHLECCGILVCKGPNLMTAKDVGKEN-RSIVLEVSAPLIVGSLPDDA 197
QY 204 AVANNYAGQVGLTASENGVEVEDKSP--YVNIIVARADNKSRAIQDFVAYQTBVE 261
DB 198 AVIPGNFALANLSPKDKLCELDLSVSKYTNLVVIRSDVSPKMKIKQKLFQSPSVQ 256

RESULT 37
US-09-489-039A-10023
Sequence 10023, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:
APPLICANT: Gary Breton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US/09/489,039A
PRIORITY FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10023
LENGTH: 306
TYPE: PR1
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10023

Query Match 15.1%; Score 209; DB 4; Length 306;
Best Local Similarity 28.7%; Pred. No. 1.4e-13;
Matches 71; Conservative 39; Mismatches 109; Indels 28; Gaps 8;

QY 38 TIKV--GWAQPEQVAEVAAGVAKKXNLVLEFENDYAMPNSAVSKGELDANAMQ 94

Tue Jun 22 11:04:04 2004

us-10-018-672-2 rat

Page 15

Db 61 TLKIHPEPAMAG-EORIIIEYINQIAPDYLGLLEAVGVODPVQADRAVAEGVAGTIYQH 119
QY 95 KEPLER--DSOEKGNLTVGNTFVYPLAGYSTKIKTLNELKQATIAVNDPSNLARA 152
Db 120 QWMLQGVVDANGFALSTVYV---FQMAFGIYSRYSVQALPENGATTVDPDQANOGQA 176
QY 153 LILEKQGLIKLKDNTLFFSTTL-DIVENPKQLVKEVDTSVAAARAIDVDLAVNNNYA 211
Db 177 LMLTORIGLISIDPAVEPRITAKLKNVGNPHQFVKEHDLTMRPALNSVDAAI---GYV 233
QY 212 GQ--VGLASENGVEDEKDSFYNNIIYARADNDSKAIQDFVAKYQDVEAEAKKQFK 269
Db 234 SDFADKVPREKGLFPAPRTFASQVIGT-----PYLSQENITVTKQAFS 280
QY 270 DGVIKGW 276
Db 281 DPTQTM 287
RESULT 38
US-09-134-000C-3687
; Sequence 3687, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3687
; LENGTH: 91
; TYPE: PRN
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3687

Query Match 7.8%; Score 108; DB 4; Length 91;
Best Local Similarity 25.8%; Pred. No. 0.00085;
Matches 23; Conservative 25; Mismatches 31; Indels 10; Gaps 2;

QY 190 DTSVAAARAIIDVDLA--VNNNNAGQVGLTASENGVEDEKDSFYNNIIYARADNDSKA 247
Db 2 DAAVATTLTYDNEBGAVALINSFAVDQGLNPKQALALEKSSPYANIIVKREDENNEN 61
QY 248 IQDFVAKYQDVEAEAKKQFKDGVIKGW 276
Db 62 VKKLIVKTLRSKEVQ-----DWITKKM 82
RESULT 39
US-09-328-352-6764
; Sequence 6764, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6764
; LENGTH: 347
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6764

Query Match 7.7%; Score 106; DB 4; Length 347;
Best Local Similarity 26.0%; Pred. No. 0.012;

Matches 74; Conservative 33; Mismatches 118; Indels 60; Gaps 14;
QY 11 ALASGIALAGCSQNSNEPAA--ISKTPAQT--IKGVWAGPEQAVAEVAGQVAKER---- 62
Db 17 SILSAVVLGGCDUTAKVPEAKQDAQAAMTKPTITIGSDWP-----GWVAMQVAIKQGWLK 72
QY 63 -YNIIVELVFNPNYAMPNSAVSGELDANAMOKHPELEKDSOEKGLNNIVY-----GN 115
Db 73 EAGLNEFKMF-DYSASLSFSAHQDPAVLVNGDVLVRS--GTCQWMLTATDYSAGN 129
QY 116 TTVYPLAGYSTKIKTLNELKQATIAVNDPSNLARALILEKQGLIKLKDNTLFFSTTL 175
Db 130 DVIIIAKEG---INTIQLDK-GKSGIV-----EKGIV---DHLLATLAL 165
QY 176 ---DIVENPKQLVKEVDTSVAAARAIDVDLAVNNNNAGQVGLTASENGVEDEKDSF- 231
Db 166 TDHNIKSNVEKLVNSATNQDPQVFNSPDISALAIVQPVAGQALKAVAGSKIIITYSKDKFG 225
QY 232 -----YNNIIYARADNDSKAI---QDFVAKYQDVEAEAKK 266
Db 226 LIYDTLITVNMSHLTAHQEEMKILIQWDKTVKYNPDATHADAVK 270

RESULT 40
US-09-134-000C-5518
; Sequence 5518, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5518
; LENGTH: 794
; TYPE: PRN
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5518

Query Match 7.6%; Score 105; DB 4; Length 794;
Best Local Similarity 22.5%; Pred. No. 0.059;
Matches 64; Conservative 51; Mismatches 131; Indels 38; Gaps 10;

QY 1 MNFGKINGICALASGIALAGCSQNSNEPAAISKTAQITKGVWAGPEQAVAEVAGQVAK 60
Db 387 VNNNNILRBJNGKRAQVKSNDAPDQOSNANASWMAAQF-GVGSVAFPKQ-LENGKNNFLK 444
QY 61 EKYNTLVELVFNPNYAMPNSAV---SKGELDANAMOKHPELEKDSOEKGLNNIVY----- 111
Db 445 DNYS-----LLEGNYPEKEHEDVVLIVDSNNTNINAKLNGFQVKNQKIASDYGTMK 500
QY 112 TVGNTFVYPLAGYSTKIKTLNELKQATIAVNDPSNLARALILEKQGLIKLKDNTLFFSTTL 169
Db 501 LANNNAFY-----TKLPTGNFIPQDLOAVYDNEBNE-----LTISGILRIKDSSTN 549
QY 170 LFTSTLDIVENPKQLVKEVDTSVAAARAIDVDLAVNNNNAGQVGLTASENGVEDEKDSF 229
Db 550 LLAPGIAVSDALSTNMIAKNTSDIVKQASNTVMTNNE---TLDASAKENLSTLYGAN 606
QY 230 SPYNNIIYARADNDSKAIQDFVAKYQDVEAEAKKQFKDGVY 273
Db 607 EIPSSIMTYPNDFKSKIDLDLDAVYNG-----KDKKQI 643

Search completed: June 16, 2004, 11:09:17
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:09:22 ; Search time 17 Seconds

(without alignments)
845.374 Million cell updates/sec

Title: US-10-018-672-2

Perfect score: 276
Sequence: 1 NMFGKINGICALASGIALAG.....TDEVEAKKQPKGVIKGW 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34	12.3	277	1	PLPB_PASHA	O08668 pasteurilla
2	19	6.9	273	1	METQ_HAEN	P11728 haemophilus
3	12	4.3	263	1	PLPB_PASHA	O08670 pasteurilla
4	12	4.3	276	1	METQ_PASHU	Q9C835 pasteurilla
5	11	4.0	271	1	METQ_ECO57	Q88V99 escherichia
6	11	4.0	271	1	METQ_SALTI	Q82992 salmonella
7	11	4.0	271	1	METQ_SALTI	Q82991 salmonella
8	9	3.3	125	1	RPA3_YEAST	P25329 saccharomyc
9	9	3.3	269	1	METQ_VIBCH	Q9K177 vibrio chol
10	9	3.3	391	1	MUCB_HUMAN	P04220 homo sapien
11	8	2.9	74	1	YD97_CAMKE	Q9PM70 campylobact
12	8	2.9	271	1	METQ_ECOLI	P28635 escherichia
13	8	2.9	837	1	SM4G_MOUSE	Q9WH17 mus musculu
14	8	2.9	838	1	SM4G_MOUSE	Q9WH17 mus musculu
15	8	2.9	1220	1	LBNI_MOUSE	Q9WH17 mus musculu
16	7	2.5	103	1	IHEB_XANAC	Q8PK32 xanthomonas
17	7	2.5	103	1	IHEB_XANCP	Q8PK32 xanthomonas
18	7	2.5	104	1	IHEB_XYLPF	Q8PK38 xylella fas
19	7	2.5	104	1	IHEB_XYLPF	Q8PK38 xylella fas
20	7	2.5	141	1	PSPI_PSEIC	Q81704 pseudopust
21	7	2.5	193	1	COAT_AFLSA	Q81704 pseudopust
22	7	2.5	233	1	MRKB_KLEPN	P21646 klebsiella
23	7	2.5	240	1	KGUA_MYCGE	P47353 mycoplasma
24	7	2.5	268	1	CLCR_RAT	P55091 rattus norv
25	7	2.5	268	1	SAPF_ECOLI	P26638 escherichia
26	7	2.5	268	1	SAPF_SALTY	P26638 salmonella
27	7	2.5	269	1	SAPF_HAEN	P25289 haemophilus
28	7	2.5	271	1	METQ_YEPE	Q08H40 yersinia pe
29	7	2.5	276	1	PLPB_PASHA	Q08669 pasteurilla
30	7	2.5	291	1	NAK2_ECOL6	Q8FDU8 escherichia
31	7	2.5	297	1	ADT1_BOVIN	P02722 bos taurus
32	7	2.5	297	1	YD67_METUA	Q58762 methanococ
33	7	2.5	298	1	ADT2_HUMAN	P05141 homo sapien
34	7	2.5	298	1	ADT2_HUMAN	P05141 homo sapien

107	6	2.2	130	1	HIS3_PSEBM	Q87uy9 pseudomonas	180	6	2.2	201	1	LEUD_METEX	Q87p38 methyllobact
108	6	2.2	134	1	HIS3_PSEBM	Q87uy9 pseudomonas	181	6	2.2	202	1	BIN3_STAU	P03348 staphylococ
109	6	2.2	134	1	Y652_MENTA	Q58068 methanococ	182	6	2.2	203	1	TRMB_MYCPU	Q68r44 mycoplasma
110	6	2.2	137	1	COF1_DICDI	P54706 dictyostell	183	6	2.2	204	1	RUVA_HABIN	P44632 haemophilus
111	6	2.2	138	1	RBFA_PSEBM	Q87wq6 pseudomonas	184	6	2.2	204	1	RUVA_VIBPA	Q67q48 vibrio para
112	6	2.2	143	1	PSAG_HORVU	Q00327 hordium vil	185	6	2.2	204	1	RUVA_VIBU	Q68d43 vibrio vuln
113	6	2.2	143	1	R191_ARATH	Q99346 arabidopsis	186	6	2.2	206	1	THIE_OCBIR	Q62j32 yerersinia pe
114	6	2.2	143	1	R192_ARATH	Q91f30 arabidopsis	187	6	2.2	207	1	KGUA_YERPE	Q74175 schizosacch
115	6	2.2	144	1	USPF_ECOL6	Q87np3 escherichia	188	6	2.2	208	1	RLI3_SCHPO	P10135 mycoplasma
116	6	2.2	144	1	USPF_SALTY	Q87b6 salmonella	189	6	2.2	208	1	RL4_MYCAA	P16976 zea mays (m
117	6	2.2	146	1	GLB1_ANABR	P02212 anadara bro	190	6	2.2	208	1	VPTI_MAIZE	Q6nnr4 corynebacte
118	6	2.2	146	1	GLB1_SCATN	P02213 scapharca i	191	6	2.2	209	1	RNH2_CONGL	Q63578 l putative
119	6	2.2	146	1	GLB1_ANATR	P14394 canadara tra	192	6	2.2	210	1	ALKH_TREPA	P75599 mycoplasma
120	6	2.2	147	1	RS5_CAMOE	Q9p1y8 campylobact	193	6	2.2	210	1	TRMB_MYCCE	P75599 mycoplasma
121	6	2.2	148	1	RL9_LISMO	Q87ar2 listeria mo	194	6	2.2	211	1	HET3_RADMG	Q6u6r1 radianthus
122	6	2.2	148	1	RL9_PSEBK	Q88d1 pseudomonas	195	6	2.2	211	1	MDCG_PSEBM	Q87v59 pseudomonas
123	6	2.2	149	1	RL9_PSEBM	Q87v59 pseudomonas	196	6	2.2	212	1	FLBE_CAUCR	Q04953 caulobacter
124	6	2.2	149	1	NABE_ECOLI	P33393 escherichia	197	6	2.2	213	1	CYSE_STEAM	Q99w74 staphylococ
125	6	2.2	149	1	YID1_ECOLI	P31446 escherichia	198	6	2.2	215	1	PSB6_YEAST	P38524 saccharomyc
126	6	2.2	151	1	DTD_IACLA	Q9Cj92 lactococcus	199	6	2.2	215	1	YKJ7_YEAST	P28757 saccharomyc
127	6	2.2	151	1	RS9_MYCTU	Q06259 mycobacteri	200	6	2.2	218	1	TRKA_METUA	Q85055 methanococ
128	6	2.2	153	1	NDK3_SPIOL	P81766 spinacia ol	201	6	2.2	218	1	Y522_METUA	Q37942 methanococ
129	6	2.2	153	1	RS9_MYCLE	P40828 mycobacteri	202	6	2.2	220	1	Y064_ARCFU	Q30112 archaeoglob
130	6	2.2	155	1	CYNS_PSEBM	Q885a6 pseudomonas	203	6	2.2	222	1	SFSA_THRAC	Q3hk14 thermoplas
131	6	2.2	155	1	REG1_PYRAB	Q9v2m0 pyrococcus	204	6	2.2	225	1	ATP6_LOCOM	P45569 locusta mig
132	6	2.2	155	1	REG1_PYRHO	Q57818 pyrococcus	205	6	2.2	226	1	HAMI_CHUTE	Q6Kff6 chlorobium
133	6	2.2	156	1	BCCP_PSEAE	P37799 pseudomonas	206	6	2.2	226	1	INH_BP74	P88038 bacterioph
134	6	2.2	157	1	GREB_SALTY	Q8z317 salmonella	207	6	2.2	228	1	YI41_ARCFU	Q8141 archaeoglob
135	6	2.2	157	1	GREB_SALTY	Q8z1j2 salmonella	208	6	2.2	229	1	NEP1_PYRHO	Q50081 pyrococcus
136	6	2.2	158	1	GREB_XANCP	Q8p1d9 xanthomonas	209	6	2.2	229	1	Y767_CABEL	Q10083 caenorhabd
137	6	2.2	158	1	GREB_XANCP	Q914d3 xanthomonas	210	6	2.2	231	1	OVAX_CHICK	Q34332 gallus gall
138	6	2.2	158	1	GREB_ECOL6	Q8x723 escherichia	211	6	2.2	232	1	YEPH_ECOLI	P16551 escherichia
139	6	2.2	158	1	GREB_ECOLI	P30128 escherichia	212	6	2.2	232	1	YEPH_ECOLI	Q92920 streptococ
140	6	2.2	160	1	TIMH_CABEL	Q21655 caenorhabd	213	6	2.2	234	1	NAGB_STIRP	P84429 agrobacteri
141	6	2.2	160	1	COTF_BACSU	P2361 bacillus su	214	6	2.2	239	1	SFSA_AGR15	Q05550 salmonella
142	6	2.2	161	1	YB83_MENTU	Q58585 methanococ	215	6	2.2	241	1	CBID_SALTY	P05525 bacillus ca
143	6	2.2	165	1	VEAE_BPP22	Q03547 bacterioph	216	6	2.2	242	1	TRPD_BACCA	Q8p663 xanthomonas
144	6	2.2	167	1	TC25_YEAST	P44277 haemophilus	217	6	2.2	242	1	YU27_XANCP	Q8p663 xanthomonas
145	6	2.2	167	1	TC25_YEAST	P17477 saccharomyc	218	6	2.2	242	1	YU51_XANCP	Q8p663 xanthomonas
146	6	2.2	173	1	YR15_CABEL	Q09355 caenorhabd	219	6	2.2	243	1	YRHO_SALTY	Q8xep9 salmonella
147	6	2.2	175	1	OMLA_PSEFL	Q68564 pseudomonas	220	6	2.2	245	1	PHOS_FELCA	Q66374 colletoctic
148	6	2.2	176	1	RIMM_THEMA	Q9x1q4 thermocoga	221	6	2.2	245	1	PYRF_YERPE	P81666 felis silve
149	6	2.2	176	1	SSB_HABHU	P59930 haemophilus	222	6	2.2	245	1	YABS_BACSU	P75561 bacillus pe
150	6	2.2	178	1	GRPE_RICPR	Q9zct4 rickettsia	223	6	2.2	245	1	YRHO_ECOLI	Q8x83 escherichia
151	6	2.2	178	1	YB16_SCHPO	Q10195 schizosacch	224	6	2.2	245	1	YRHO_ECOLI	P77488 escherichia
152	6	2.2	180	1	YB16_BUCAP	Q8x831 buchnera ap	225	6	2.2	246	1	RL7A_SCHPO	P17937 schizosacch
153	6	2.2	180	1	YCBU_ECOLI	Q8d1g8 synchococc	226	6	2.2	249	1	LACR_STIAM	Q99873 staphylococ
154	6	2.2	182	1	IF3_SYNEL	P71578 mycobacteri	227	6	2.2	251	1	SODE_MOUSE	P16644 staphylococ
155	6	2.2	182	1	ORN_COXBU	Q83c93 coxiella bu	228	6	2.2	251	1	NAGB_STIAM	Q99w74 mus muscu
156	6	2.2	183	1	SPP2_YEAST	Q002521 saccharomyc	229	6	2.2	252	1	VG12_BPM2	Q84265 mycobacteri
157	6	2.2	185	1	T2_MOUSE	Q06666 mus muscu	230	6	2.2	254	1	HIS6_NEIMA	Q91v55 neisseria m
158	6	2.2	185	1	ENGB_MYCPU	Q984x1 mycoplasma	231	6	2.2	255	1	HIS6_NEIMA	Q83014 neisseria m
159	6	2.2	186	1	ISP2_VIBPA	Q87nas vibrio para	232	6	2.2	255	1	SVUB_MYCLE	P83045 mycobacteri
160	6	2.2	187	1	Y768_AQUAE	P05014 homo sapien	233	6	2.2	255	1	KSQA_MYCCE	P87701 m dimethyla
161	6	2.2	188	1	INAA_HUMAN	P01567 homo sapien	234	6	2.2	259	1	YAZO_METUA	Q84426 methanococ
162	6	2.2	189	1	INAA_HUMAN	P01567 homo sapien	235	6	2.2	259	1	CLCA_PSEPU	P11451 pseudomonas
163	6	2.2	189	1	INAA_HUMAN	P01567 homo sapien	236	6	2.2	260	1	Y997_STRCO	Q91016 streptomyce
164	6	2.2	189	1	INAA_HUMAN	P01567 homo sapien	237	6	2.2	260	1	GLNH_BACST	P27676 bacillus st
165	6	2.2	189	1	INAA_HUMAN	P01567 homo sapien	238	6	2.2	262	1	PNP2_BACHD	Q8x808 bacillus ha
166	6	2.2	189	1	INAA_HUMAN	P01567 homo sapien	239	6	2.2	265	1	GNP1_CABEL	Q9xv72 caenorhabd
167	6	2.2	189	1	INAA_HUMAN	P01567 homo sapien	240	6	2.2	267	1	MEMB_CVACA	Q9tm10 cyandidium c
168	6	2.2	192	1	COMK_BACSU	P106322 bacillari	241	6	2.2	268	1	TR1_DROME	P36188 drosophila
169	6	2.2	192	1	YD14_MYCTU	Q97ry6 detinococcus	242	6	2.2	269	1	Y9C0_BACCA	Q81322 bacillus ce
170	6	2.2	194	1	Y667_DEIRA	Q8p1y8 xanthomonas	243	6	2.2	270	1	NLPA_ECOLI	Q8x808 escherichia
171	6	2.2	196	1	YHMF_ECOLI	P10120 escherichia	244	6	2.2	272	1	Y12F_BP74	P58220 rhizobium 1
172	6	2.2	198	1	CYCO_BACSU	Q06735 bacillus su	245	6	2.2	273	1	DABI_RHILLO	P82215 bacterioph
173	6	2.2	199	1	Y470_DEIRA	Q97ry6 detinococcus	246	6	2.2	279	1	TRPA_RHILLO	Q88cm6 rhizobium 1
174	6	2.2	199	1	Y470_DEIRA	Q97ry6 detinococcus	247	6	2.2				
175	6	2.2	199	1	Y470_DEIRA	Q97ry6 detinococcus	248	6	2.2				
176	6	2.2	199	1	Y470_DEIRA	Q97ry6 detinococcus	249	6	2.2				
177	6	2.2	199	1	Y470_DEIRA	Q97ry6 detinococcus	250	6	2.2				
178	6	2.2	199	1	Y470_DEIRA	Q97ry6 detinococcus	251	6	2.2				
179	6	2.2	200	1	RR4_GUTH	P17072 guillardia	252	6	2.2				

253	6	2.2	280	1	OXA2_BACHD	Q9kdp2 bacillus ha	326	6	2.2	335	1	RUVB_CAMTE	Q9pmt7 campylobact
254	6	2.2	284	1	EUTC_PROHL	Q7n2b3 phocorhabd	327	6	2.2	336	1	DHOW_METTA	Q859m7 methanococ
255	6	2.2	285	1	ECHC_MYCTU	Q531e3 mycobacteri	328	6	2.2	336	1	GCP_BUCAL	P57166 buchnera ap
256	6	2.2	289	1	ATP6_PROMO	P21903 propionigen	329	6	2.2	336	1	HUTG_VIBVU	Q8da19 vibrio ap
257	6	2.2	291	1	YOHO_BACSU	P54513 bacillus su	330	6	2.2	336	1	HUTG_VIBVU	P60111 vibrio valn
258	6	2.2	292	1	DHVS_PYRAE	Q8z09 pyrobaculum	331	6	2.2	336	1	RUVB_HELPJ	Q9ams7 helicobacte
259	6	2.2	293	1	REPFI_ECOLI	P27831 escherichia	332	6	2.2	337	1	RUVB_HELPJ	Q25699 helicobacte
260	6	2.2	294	1	NPLT_ECOLI	P39933 escherichia	333	6	2.2	337	1	RCGA_SUISO	Q97w04 sulfobobus
261	6	2.2	294	1	PTA_THEMA	Q9x014 thermocoga	334	6	2.2	338	1	MDHM_HUMAN	P04926 homo sapien
262	6	2.2	296	1	NADC_RHORI	P77938 rhodospirill	335	6	2.2	338	1	MDHM_MOUSE	P08242 mus musculu
263	6	2.2	296	1	NIFH_PLEBO	Q00240 plectonema	336	6	2.2	338	1	MDHM_PIG	P00346 sus scrofa
264	6	2.2	298	1	DAZL_MOUSE	Q64368 mus musculu	337	6	2.2	338	1	MDHM_RAT	P04636 rattus norv
265	6	2.2	298	1	MMSB_PSEAE	P28811 pseudomonas	338	6	2.2	339	1	MOAA_BACHD	Q9k9w9 bacillus ha
266	6	2.2	299	1	CSNS_SCHPO	Q94454 schizosacch	339	6	2.2	339	1	ADT_GILKE	P31692 chlorella k
267	6	2.2	299	1	NANA_RHIME	Q92w00 rhizobium m	340	6	2.2	340	1	CEAA_BACUT	Q45729 bacillus th
268	6	2.2	299	1	SUCD_BACSU	P80865 bacillus su	341	6	2.2	340	1	COND_NEUCR	P14663 neurospora
269	6	2.2	300	1	ARGB_METUA	Q60382 methanococc	342	6	2.2	341	1	Y534_TREPA	O83545 treponema p
270	6	2.2	300	1	PSD_FUSUN	Q819f2 fusobacteri	343	6	2.2	342	1	ISIA_SYNP7	P15347 synechococc
271	6	2.2	302	1	RBGR_CHEVI	P25544 chromacium	344	6	2.2	342	1	ISIA_SYNP3	O55274 synechocyst
272	6	2.2	302	1	SUCD_STPAAM	Q9nm44 staphylococ	345	6	2.2	344	1	HRCA_STREN	O544r2 streptococc
273	6	2.2	302	1	SUCD_STPAAM	Q9nm01 staphylococ	346	6	2.2	344	1	LKPD_SYNP3	O56612 synechocyst
274	6	2.2	302	1	SUCD_STPAEP	Q8c0b4 staphylococ	347	6	2.2	345	1	UI94_HCVYA	P16800 human cytom
275	6	2.2	303	1	OSTP_PIG	P14267 sus scrofa	348	6	2.2	346	1	FMIR_GORCO	P79176 gorilla gor
276	6	2.2	303	1	Y367_RICPR	Q9z4g2 ticketstia	349	6	2.2	346	1	FMIR_MACMU	P79189 macaca mula
277	6	2.2	304	1	NODI_RHIS3	P72335 rhizobium s	350	6	2.2	346	1	FMIR_PANTR	P79241 pan troglod
278	6	2.2	305	1	E2BA_HUMAN	Q14232 homo sapien	351	6	2.2	346	1	FMIR_PONPY	P79235 pongo pygma
279	6	2.2	305	1	NODI_BRASS	Q92313 bradyrhizob	352	6	2.2	347	1	HPT_MOUSE	O61646 mus musculu
280	6	2.2	306	1	NODI_BRAJA	P26050 bradyrhizob	353	6	2.2	347	1	HPT_MOUSE	O60574 mus caroli
281	6	2.2	307	1	YDCI_ECOLI	P77171 escherichia	354	6	2.2	347	1	HPT_MOUSE	Q62558 mus saxicol
282	6	2.2	308	1	MEIR_ECOLI	Q8x610 escherichia	355	6	2.2	347	1	HPT_PIG	Q86p57 sus scrofa
283	6	2.2	308	1	MEIR_ECOLI	P07623 escherichia	356	6	2.2	347	1	HPT_RAT	P06866 rattus norv
284	6	2.2	308	1	MEIR_SALTI	Q8z1w1 salmonella	357	6	2.2	347	1	SELD_ECOLI6	O8f9y3 escherichia
285	6	2.2	308	1	MEIR_SALTI	P37413 salmonella	358	6	2.2	347	1	SELD_ECOLI	P145456 escherichia
286	6	2.2	308	1	RNH3_LISIN	Q92c10 listeria in	359	6	2.2	347	1	SELD_SALTI	O8x6f4 salmonella
287	6	2.2	308	1	RNH3_LISIN	Q8y0p5 listeria in	360	6	2.2	347	1	SELD_SALTI	O8x6p5 salmonella
288	6	2.2	309	1	ARCL_MYCPN	P78030 mycoplasma	361	6	2.2	347	1	SELD_SHITL	P58393 shigella fl
289	6	2.2	309	1	CRFS_MOUSE	Q9c226 mus musculu	362	6	2.2	348	1	HPTR_HUMAN	P00739 homo sapien
290	6	2.2	309	1	KHSE_SALTY	Q8x2p5 salmonella	363	6	2.2	348	1	SELD_HAEIN	P43911 haemophilus
291	6	2.2	310	1	PPAC_STRMU	O68579 streptococc	364	6	2.2	348	1	XYLD_RHIOI	Q96d10 rhizobium l
292	6	2.2	310	1	TF2B_METH	O26971 methanobact	365	6	2.2	349	1	ASFP_VULVU	P53533 vulpes vulp
293	6	2.2	311	1	HPRK_MYCGE	P47331 mycoplasma	366	6	2.2	350	1	FMIR_HUMAN	P21465 homo sapien
294	6	2.2	312	1	CELF_PVKCA	O85230 pseudorabie	367	6	2.2	350	1	MRP_DEIRA	O97462 deinozocous
295	6	2.2	313	1	HEM3_SALTY	Q916q2 salmonella	368	6	2.2	352	1	FMIX_ECOLI	P10122 escherichia
296	6	2.2	313	1	MOAA_SULTO	Q95711 sulfolobus	369	6	2.2	354	1	BCA4_ARATH	O91e06 arabidopsis
297	6	2.2	313	1	YDIR_BACSU	Q50225 bacillus su	370	6	2.2	355	1	NDI1_RHIME	O52618 rhizobium m
298	6	2.2	314	1	FATD_VIBAN	P37738 vibrio angu	371	6	2.2	356	1	GBA2_USTWA	P87033 ustilla go ma
299	6	2.2	314	1	TYSY_CANAL	P12461 candida alb	372	6	2.2	356	1	LAR2_RAT	Q924u0 rattus norv
300	6	2.2	316	1	CZCD_ALCOB	P13512 alcaligenes	373	6	2.2	359	1	DP04_CLOPE	O8x583 homo sapien
301	6	2.2	316	1	CZCD_ALCOB	P94178 alcaligenes	374	6	2.2	359	1	YD56_AGR75	P07585 homo sapien
302	6	2.2	316	1	RAM2_YEAST	P29703 s protein f	375	6	2.2	359	1	YD56_AGR75	Q8f8p1 agrobacteri
303	6	2.2	316	1	Y034_TREPA	O83037 treponema p	376	6	2.2	359	1	YD56_AGR75	Q93556 caenorhabd
304	6	2.2	318	1	ATRA_ARATH	P92953 arabidopsis	377	6	2.2	360	1	LAR2_MOUSE	Q93j19 mus musculu
305	6	2.2	318	1	KHSE_VIBCB	Q9x0k4 vibrio chol	378	6	2.2	360	1	PGS2_BOVIN	P21793 bos taurus
306	6	2.2	321	1	YB28_SYNP3	P72655 synchocyst	379	6	2.2	360	1	PGS2_CANPA	Q29399 canis famli
307	6	2.2	321	1	YDGT_SCHPO	Q10494 schizosacch	380	6	2.2	360	1	PGS2_PIG	Q28888 oryzocolagus
308	6	2.2	322	1	RADA_METVO	O73948 methanococ	381	6	2.2	360	1	PGS2_RABIT	Q28888 oryzocolagus
309	6	2.2	322	1	TPIC_SPIOL	P48496 spirinacia ol	382	6	2.2	361	1	AROB_BACHD	O9x0cb bacillus ha
310	6	2.2	322	1	Y711_METUA	Q58j21 methanococ	383	6	2.2	362	1	SERC_XYLFA	Q9pb19 xyella fas
311	6	2.2	323	1	OXSJ_HUMAN	O8n9j8 homo sapien	384	6	2.2	362	1	SERC_XYLFA	Q8tbn0 xyella fas
312	6	2.2	325	1	CYSK_SOLTU	O81154 solanum tub	385	6	2.2	362	1	PELB_ASPEL	P41750 aspergillus
313	6	2.2	326	1	TF2B_ARCFU	O28970 archaeoglob	386	6	2.2	366	1	PUR7_VIBCH	Q9k9t5 vibrio chol
314	6	2.2	328	1	YIAO_HAEIN	P44992 haemophilus	387	6	2.2	368	1	DRDO_DICDI	P07670 dictyosteli
315	6	2.2	329	1	HPT_CANPA	P19006 canis famli	388	6	2.2	369	1	G6P2_CAPAN	P80042 capsicum an
316	6	2.2	329	1	SUBI_SALTY	P02906 salmonella	389	6	2.2	369	1	RADA_CERSY	O33748 cenarchaeum
317	6	2.2	330	1	ODER_BACSU	P37940 bacillus su	390	6	2.2	370	1	GIDA_BACST	P32816 bacillus st
318	6	2.2	332	1	AAPR_WHEAT	O20266 trititum ae	391	6	2.2	370	1	TRM1_SULTO	P37199 sulfolobus
319	6	2.2	332	1	GPDA_CLOAB	O971d6 clostridium	392	6	2.2	374	1	KAPR_APICA	P31319 alvaysa cal
320	6	2.2	332	1	MYOD_DROME	P22816 drosophila	393	6	2.2	377	1	SEN2_YEAST	P16658 saccharomyc
321	6	2.2	333	1	PLSX_BACSU	P71018 bacillus su	394	6	2.2	377	1	ME24_ASPEL	P46073 aspergillus
322	6	2.2	334	1	ILVC_STPAEP	O8c9b5 staphylococ	395	6	2.2	381	1	T10B_MOUSE	Q9qzma mus musculu
323	6	2.2	334	1	OPPF_ECOLI	P77737 escherichia	396	6	2.2	381	1	OVALL_COTUY	P19104 coturnix co
324	6	2.2	334	1	OPPF_SALTY	P08007 salmonella	397	6	2.2	382	1	OPGG_PBSBY	P20400 pseudomonas
325	6	2.2	335	1	NDI2_RHIME	Q8ghn6 rhizobium m	398	6	2.2	384	1		

399	6	2.2	384	1	PKR2_HUMAN	Q9NFJ6	homo sapien	472	6	2.2	451	1	ERG8_YEAST	P24521	saccharomyc
400	6	2.2	385	1	DIAC_HUMAN	Q01459	homo sapien	473	6	2.2	453	1	BRF3_YEAST	P24521	saccharomyc
401	6	2.2	385	1	OVAL_CHICK	P01012	gallus gall	474	6	2.2	457	1	MBL_ECOLI	P37773	escherichia
402	6	2.2	387	1	ENG6_VIBCH	Q9KX18	vibrio chol	475	6	2.2	457	1	YC47_SYNY3	P45411	synecocyst
403	6	2.2	387	1	NMT_DROME	Q61613	drosophila	476	6	2.2	461	1	NCB1_HUMAN	Q02818	homo sapien
404	6	2.2	388	1	AMPK_VEREN	P45600	versinia en	477	6	2.2	462	1	ALB3_ARATH	Q91B94	arabidopsis
405	6	2.2	388	1	MSOX_STRESS	P40854	streptomyce	478	6	2.2	463	1	NAOX_METUA	Q80655	methanococ
406	6	2.2	388	1	OVAT_CHICK	P01014	gallus gall	479	6	2.2	463	1	SYG1_RICAR	Q82299	wiggleswort
407	6	2.2	390	1	DXR_FUSRN	Q81622	fusobacteri	480	6	2.2	463	1	SYG1_STAM	Q99111	straphyleocc
408	6	2.2	391	1	DXR_LEPIN	Q81646	leptospira	481	6	2.2	463	1	SYG1_STAM	Q99111	straphyleocc
409	6	2.2	391	1	XRCL_THETN	Q81646	thermoanaer	482	6	2.2	464	1	SYG1_CLOAB	Q99111	straphyleocc
410	6	2.2	392	1	DXR_CORGL	Q81646	thermoanaer	483	6	2.2	464	1	6PGD_SALTY	Q99111	straphyleocc
411	6	2.2	392	1	GALI_CANFA	Q81646	thermoanaer	484	6	2.2	464	1	EFIA_HYDAT	P15554	hydra atten
412	6	2.2	392	1	MP12_AMBAR	Q81646	thermoanaer	485	6	2.2	469	1	C39A_HUMAN	Q91515	homo sapien
413	6	2.2	394	1	PGK_BACHD	Q81646	thermoanaer	486	6	2.2	470	1	SUP1_ECOLI	P26176	rhodobacter
414	6	2.2	395	1	CYSK_SCHPO	Q81646	thermoanaer	487	6	2.2	472	1	PECK_TRYBB	P37375	trypanosoma
415	6	2.2	396	1	ATPB_ADIPA	Q81646	thermoanaer	488	6	2.2	472	1	VACA_BACSU	P37375	trypanosoma
416	6	2.2	396	1	DEOB_CLOPE	Q81646	thermoanaer	489	6	2.2	473	1	DNPE_MOUSE	P37375	trypanosoma
417	6	2.2	396	1	MP11_AMBAR	Q81646	thermoanaer	490	6	2.2	473	1	REBM_SALMU	Q00473	salmonella
418	6	2.2	397	1	DXR_HAETN	Q81646	thermoanaer	491	6	2.2	473	1	NIEB_ANANZ	Q38813	anabaena az
419	6	2.2	397	1	MP13_AMBAR	Q81646	thermoanaer	492	6	2.2	475	1	NIEB_ANANZ	Q38813	anabaena az
420	6	2.2	398	1	MP12_AMBAR	Q81646	thermoanaer	493	6	2.2	475	1	NIEB_ANANZ	Q38813	anabaena az
421	6	2.2	398	1	PGK_LACLA	Q81646	thermoanaer	494	6	2.2	477	1	YFDM_RHOCA	P26176	rhodobacter
422	6	2.2	400	1	AATA_RHIME	Q81646	thermoanaer	495	6	2.2	479	1	ERS2_CHICK	P26176	rhodobacter
423	6	2.2	402	1	CPT2_BRAOL	Q81646	thermoanaer	496	6	2.2	482	1	VS12_TRYBB	P26176	rhodobacter
424	6	2.2	402	1	PGK_CHLEP	Q81646	thermoanaer	497	6	2.2	482	1	VS12_TRYBB	P26176	rhodobacter
425	6	2.2	405	1	ASSY_PSEAE	Q81646	thermoanaer	498	6	2.2	483	1	SYP_MYCPN	P26176	rhodobacter
426	6	2.2	405	1	ASSY_PSEAE	Q81646	thermoanaer	499	6	2.2	483	1	VS12_TRYBB	P26176	rhodobacter
427	6	2.2	405	1	ASSY_PSEAE	Q81646	thermoanaer	500	6	2.2	485	1	VS12_TRYBB	P26176	rhodobacter
428	6	2.2	405	1	HOW_DROME	Q81646	thermoanaer	501	6	2.2	485	1	VS12_TRYBB	P26176	rhodobacter
429	6	2.2	405	1	PRSB_YEAST	Q81646	thermoanaer	502	6	2.2	486	1	VS12_TRYBB	P26176	rhodobacter
430	6	2.2	406	1	ASSY_CAMJE	Q81646	thermoanaer	503	6	2.2	486	1	VS12_TRYBB	P26176	rhodobacter
431	6	2.2	406	1	HPT_HUMAN	Q81646	thermoanaer	504	6	2.2	486	1	VS12_TRYBB	P26176	rhodobacter
432	6	2.2	408	1	YF1N_ECOLI	Q81646	thermoanaer	505	6	2.2	487	1	WRK1_ARATH	Q38137	arabidopsis
433	6	2.2	412	1	YAG7_SCHPO	Q81646	thermoanaer	506	6	2.2	490	1	YHFE_ECOLI	P37375	trypanosoma
434	6	2.2	413	1	APGM_SUTJO	Q81646	thermoanaer	507	6	2.2	490	1	YHFE_ECOLI	P37375	trypanosoma
435	6	2.2	414	1	CUTS_STRCO	Q81646	thermoanaer	508	6	2.2	493	1	CPE1_YEAST	P37375	trypanosoma
436	6	2.2	414	1	K193_HUMAN	Q81646	thermoanaer	509	6	2.2	494	1	MORE_BACSU	Q03552	bacillus su
437	6	2.2	414	1	K193_MOUSE	Q81646	thermoanaer	510	6	2.2	495	1	IMDH_METUA	P37375	trypanosoma
438	6	2.2	414	1	PROA_XANCP	Q81646	thermoanaer	511	6	2.2	496	1	IMDH_METUA	P37375	trypanosoma
439	6	2.2	415	1	NMTB_MYCBO	Q81646	thermoanaer	512	6	2.2	498	1	DHAM_LEITA	Q35415	leishmania
440	6	2.2	417	1	ARPA_HUMAN	Q81646	thermoanaer	513	6	2.2	498	1	DHAM_LEITA	Q35415	leishmania
441	6	2.2	417	1	GLAI_PSEAE	Q81646	thermoanaer	514	6	2.2	500	1	LYCB_TYCES	Q35787	lycoperisico
442	6	2.2	417	1	GLAI_PSEAE	Q81646	thermoanaer	515	6	2.2	500	1	LYCB_TYCES	Q35787	lycoperisico
443	6	2.2	418	1	GLAI_PSEAE	Q81646	thermoanaer	516	6	2.2	500	1	LYCB_TYCES	Q35787	lycoperisico
444	6	2.2	418	1	NER3_MOUSE	Q81646	thermoanaer	517	6	2.2	501	1	Y039_BORBU	Q51068	borrelia bu
445	6	2.2	419	1	LMB2_YERPE	Q81646	thermoanaer	518	6	2.2	501	1	C723_ARATH	Q51068	borrelia bu
446	6	2.2	420	1	ASSY_YEAST	Q81646	thermoanaer	519	6	2.2	501	1	C723_ARATH	Q51068	borrelia bu
447	6	2.2	420	1	GLYA_PASMU	Q81646	thermoanaer	520	6	2.2	501	1	C72P_ARATH	Q51068	borrelia bu
448	6	2.2	420	1	GLYA_STRCO	Q81646	thermoanaer	521	6	2.2	501	1	C72P_ARATH	Q51068	borrelia bu
449	6	2.2	421	1	GLYA_HABIN	Q81646	thermoanaer	522	6	2.2	501	1	LYCB_TYCES	Q35787	lycoperisico
450	6	2.2	421	1	TOLA_ECOLI	Q81646	thermoanaer	523	6	2.2	502	1	LYCB_TYCES	Q35787	lycoperisico
451	6	2.2	423	1	SNX4_YEAST	Q81646	thermoanaer	524	6	2.2	503	1	KIR3_MOUSE	Q35787	lycoperisico
452	6	2.2	426	1	GLAI_MYCBO	Q81646	thermoanaer	525	6	2.2	503	1	KIR3_MOUSE	Q35787	lycoperisico
453	6	2.2	426	1	GLAI_MYCBO	Q81646	thermoanaer	526	6	2.2	506	1	C728_ARATH	Q35787	lycoperisico
454	6	2.2	426	1	GLAI_MYCBO	Q81646	thermoanaer	527	6	2.2	506	1	C728_ARATH	Q35787	lycoperisico
455	6	2.2	428	1	NMTB_MYCBO	Q81646	thermoanaer	528	6	2.2	508	1	COX1_BLAGE	Q35787	lycoperisico
456	6	2.2	429	1	GUNA_BUFI	Q81646	thermoanaer	529	6	2.2	508	1	COX1_BLAGE	Q35787	lycoperisico
457	6	2.2	430	1	Y587_AQUAE	Q81646	thermoanaer	530	6	2.2	509	1	AVR1_BOVIN	Q35787	lycoperisico
458	6	2.2	434	1	PURA_CABEL	Q81646	thermoanaer	531	6	2.2	509	1	AVR1_BOVIN	Q35787	lycoperisico
459	6	2.2	435	1	POTE_HABIN	Q81646	thermoanaer	532	6	2.2	509	1	AVR1_MOUSE	Q35787	lycoperisico
460	6	2.2	435	1	Y413_ARATH	Q81646	thermoanaer	533	6	2.2	509	1	AVR1_MOUSE	Q35787	lycoperisico
461	6	2.2	435	1	Y326_METUA	Q81646	thermoanaer	534	6	2.2	509	1	SVH_HUMAN	Q35787	lycoperisico
462	6	2.2	437	1	IF2G_METUA	Q81646	thermoanaer	535	6	2.2	509	1	SVH_HUMAN	Q35787	lycoperisico
463	6	2.2	437	1	NCCC_ALCXX	Q81646	thermoanaer	536	6	2.2	509	1	SVH_HUMAN	Q35787	lycoperisico
464	6	2.2	439	1	YD75_MYCTU	Q81646	thermoanaer	537	6	2.2	516	1	VP6_RDVA	Q35787	lycoperisico
465	6	2.2	440	1	PPAS_MOUSE	Q81646	thermoanaer	538	6	2.2	516	1	VP6_RDVA	Q35787	lycoperisico
466	6	2.2	443	1	MCRY_METPE	Q81646	thermoanaer	539	6	2.2	518	1	ALGI_RHIME	Q35787	lycoperisico
467	6	2.2	443	1	PYRD_SCHPO	Q81646	thermoanaer	540	6	2.2	518	1	ALGI_RHIME	Q35787	lycoperisico
468	6	2.2	445	1	UBA1_CANFA	Q81646	thermoanaer	541	6	2.2	519	1	RNFC_RHOCA	Q35787	lycoperisico
469	6	2.2	447	1	ANCA_CLOTM	Q81646	thermoanaer	542	6	2.2	520	1	RNFC_RHOCA	Q35787	lycoperisico
470	6	2.2	447	1	DHE4_CORF	Q81646	thermoanaer	543	6	2.2	521	1	NTEK_AZOB	Q35787	lycoperisico
471	6	2.2	447	1	DHE4_CORGL	Q81646	thermoanaer	544	6	2.2	524	1	Y018_METUA	Q35787	lycoperisico
			448	1	MAEN_BACSU	Q05256	bacillus su				525	1	GXI1_HUMGT	P15828	humicola gr

545	6	2.2	525	1	ORPG_PSEAE	Q9nuas	pseudomonas	618	6	2.2	672	1	YO30_VIBVY	Q7mtas	vibrio vuln
546	6	2.2	529	1	YBAUF_ECOLI	P39j16	escherichia	619	6	2.2	673	1	TF2B_METUA	Q58192	methanococc
547	6	2.2	531	1	YR6G_SULTO	Q97669	sulfococcus	620	6	2.2	682	1	RECG_BACSU	Q34942	bacillus su
548	6	2.2	533	1	MUTL1_THEAO	P96082	thermus agu	621	6	2.2	688	1	EPFG_CLOPE	Q88651	clostridium
549	6	2.2	542	1	THS_METUA	Q58405	methanococ	622	6	2.2	689	1	KFLB_RAT	Q88658	rattus norv
550	6	2.2	543	1	CH60_MYCPN	P78012	mycoplasma	623	6	2.2	694	1	HAAI_YEAST	Q12753	saccharomyc
551	6	2.2	543	1	THMD_MYCPN	P75436	mycoplasma	624	6	2.2	697	1	EPFG_RICBE	Q8KbD7	ricicetusa
552	6	2.2	544	1	FLCK_BUCAP	Q689X0	buchnera ap	625	6	2.2	697	1	EPFG_RICRH	Q8KbD7	ricicetusa
553	6	2.2	544	1	GSNC_SOYBN	P48640	glycine max	626	6	2.2	699	1	EPFG_RICPC	Q92193	ricicetusa
554	6	2.2	545	1	T290_ARATH	Q04212	arabidopsis	627	6	2.2	699	1	EPFG_RICPE	Q8Kra8	ricicetusa
555	6	2.2	547	1	MUTL1_THEHT	Q9ra54	thermus the	628	6	2.2	699	1	EPFG_RICHE	Q8KbD8	ricicetusa
556	6	2.2	547	1	CH60_LEGPN	P26878	legionella	629	6	2.2	699	1	EPFG_RICMO	Q8KbD6	ricicetusa
557	6	2.2	547	1	IF37_MOUSE	Q70194	mus musculu	630	6	2.2	699	1	EPFG_RICPA	Q8KbD9	ricicetusa
558	6	2.2	548	1	IF37_HUMAN	O15371	homo sapien	631	6	2.2	699	1	EPFG_RICPR	P41084	ricicetusa
559	6	2.2	549	1	Y4KD_RHISN	P55524	rhizobium s	632	6	2.2	699	1	EPFG_RICRI	Q8KbC1	ricicetusa
560	6	2.2	553	1	YIDE_SALTI	Q84217	salmonella	633	6	2.2	699	1	EPFG_RICSI	Q8KbD8	ricicetusa
561	6	2.2	553	1	YIDE_SALTI	Q84217	salmonella	634	6	2.2	699	1	EPFG_RICST	Q8KbD8	ricicetusa
562	6	2.2	554	1	Y514_SYNT3	Q55470	synechocyst	635	6	2.2	702	1	DDX4_MOUSE	Q61496	mus musculu
563	6	2.2	555	1	PGMU_ASPOR	P57179	aspergillus	636	6	2.2	702	1	DDX4_MOUSE	Q61496	mus musculu
564	6	2.2	556	1	HCP_THIPE	P96095	thiobacilli	637	6	2.2	704	1	PURU_PPRFU	Q95ano	oryza sativ
565	6	2.2	556	1	SYO_VIBCH	Q9Kba5	vibrio chol	638	6	2.2	710	1	COG2_DROME	Q8u911	pyrococcus
566	6	2.2	559	1	HUTU_BACD	Q9Kba5	bacillus ha	639	6	2.2	710	1	EPFG_BUCBP	Q9v478	drosophila
567	6	2.2	559	1	HUTU_BACD	Q9Kba5	bacillus ha	640	6	2.2	710	1	NECB_HYDAT	P59451	buchnera ap
568	6	2.2	572	1	DPY4_RAT	Q62951	rattus norv	641	6	2.2	713	1	DDX4_RAT	P29145	hydra atten
569	6	2.2	572	1	DPY2_BOVIN	O02675	bos taurus	642	6	2.2	713	1	DDX4_RAT	Q64060	rattus norv
570	6	2.2	572	1	DPY2_CHICK	Q90635	gallus galli	643	6	2.2	714	1	YFE7_YEAST	P43556	saccharomyc
571	6	2.2	572	1	DPY2_HUMAN	O16555	homo sapien	644	6	2.2	717	1	CUAB_HUMAN	Q14620	homo sapien
572	6	2.2	572	1	DPY2_MOUSE	O08553	mus musculu	645	6	2.2	721	1	PRTP_HSVJ7	P52385	human heipe
573	6	2.2	572	1	DPY2_RAT	P49442	rattus norv	646	6	2.2	724	1	DDX4_HUMAN	Q9nq10	homo sapien
574	6	2.2	572	1	DPY4_HUMAN	O14531	homo sapien	647	6	2.2	724	1	TAP1_MOUSE	P36370	mus musculu
575	6	2.2	574	1	DPY4_MOUSE	Q05098	mus musculu	648	6	2.2	725	1	FADB_ECOS7	Q84812	e fatty oxi
576	6	2.2	574	1	VGLE_HRSVA	P03420	human respi	649	6	2.2	729	1	TAP1_RAT	Q84812	e fatty oxi
577	6	2.2	574	1	VGLE_HRSVA	P12568	human respi	650	6	2.2	731	1	NGP1_HUMAN	O13823	homo sapien
578	6	2.2	575	1	IRL2_HUMAN	Q9hb29	homo sapien	651	6	2.2	738	1	MUSC_ORYZA	P12129	oryza sativ
579	6	2.2	578	1	SYD_TREPA	Q9pqk5	ureaplasma	652	6	2.2	738	1	PLAP_HUMAN	P46620	zea mays
580	6	2.2	587	1	ASB2_HUMAN	Q9eqq7	homo sapien	653	6	2.2	739	1	BCGA_PSEPL	Q59263	homo sapien
581	6	2.2	587	1	FOLC_HUMAN	Q05933	homo sapien	654	6	2.2	748	1	TAP1_HUMAN	P59931	pseudomonas
582	6	2.2	587	1	FOLC_MOUSE	P48760	mus musculu	655	6	2.2	751	1	TREA_YEAST	Q03356	homo sapien
583	6	2.2	590	1	PHO4_NEUCR	P15710	neurospora	656	6	2.2	751	1	CUBE_STRPN	P33356	saccharomyc
584	6	2.2	591	1	VATA_CHELCV	Q82236	chlamydomophi	657	6	2.2	752	1	RRL1_JAZFP	P35594	streptococc
585	6	2.2	591	1	VATA_CHLNU	Q9pk85	chlamydia m	658	6	2.2	757	1	TRP1_CRIGR	P16512	influenza a
586	6	2.2	591	1	VATA_CHLNU	Q92939	chlamydia p	659	6	2.2	760	1	FTSH_MYCTU	Q07891	cricketulsa
587	6	2.2	591	1	VATA_CHLNU	Q84310	chlamydia t	660	6	2.2	760	1	METE_MYCTE	P96942	mycobacteri
588	6	2.2	600	1	DNAX_ERVXH	Q05647	erysipeloth	661	6	2.2	761	1	TEP1_HUMAN	O05564	mycobacteri
589	6	2.2	600	1	GLMS_HALNI	Q9ht00	h glucosami	662	6	2.2	761	1	TEP1_HUMAN	P02786	homo sapien
590	6	2.2	606	1	NUSM_BALMU	P41299	balanopter	663	6	2.2	762	1	TEP2_BRAHE	O13166	homo sapien
591	6	2.2	606	1	NUSM_BALMU	P24978	balanopter	664	6	2.2	763	1	PMT4_YEAST	P46371	saccharomyc
592	6	2.2	608	1	ALBU_RABIT	P39065	oryctolagus	665	6	2.2	765	1	TRP1_MOUSE	Q62351	mus musculu
593	6	2.2	609	1	ALBU_MERIN	Q50509	meriones un	666	6	2.2	769	1	BAR1_MOUSE	O70445	mus musculu
594	6	2.2	609	1	PABX_ARATH	Q9zga9	arabidopsis	667	6	2.2	769	1	TRF1_YEAST	P38920	saccharomyc
595	6	2.2	610	1	FTMB_DICDI	P54680	dictyosteli	668	6	2.2	770	1	TRF1_FELCA	Q9my23	felis silve
596	6	2.2	611	1	YCR3_YEAST	P53531	saccharomyc	669	6	2.2	770	1	DBP4_YEAST	P20448	saccharomyc
597	6	2.2	612	1	PCD8_MOUSE	Q9z0x1	mus musculu	670	6	2.2	776	1	TRP1_CANTA	Q9g1d3	saccharomyc
598	6	2.2	612	1	PCD8_MOUSE	Q9z0x1	mus musculu	671	6	2.2	776	1	TRP1_CHICK	Q9g997	gallus galli
599	6	2.2	612	1	PCD8_MOUSE	Q9z0x1	mus musculu	672	6	2.2	778	1	TRP1_CHICK	Q9g997	gallus galli
600	6	2.2	613	1	PCD8_MOUSE	Q9z0x1	mus musculu	673	6	2.2	778	1	TRP1_CHICK	Q9g997	gallus galli
601	6	2.2	613	1	PCD8_MOUSE	Q9z0x1	mus musculu	674	6	2.2	778	1	TRP1_CHICK	Q9g997	gallus galli
602	6	2.2	615	1	PCD8_MOUSE	Q9z0x1	mus musculu	675	6	2.2	782	1	TRP1_CHICK	Q9g997	gallus galli
603	6	2.2	619	1	DNAX_CLOPE	Q892x0	azotirizobiu	676	6	2.2	784	1	TRP1_CHICK	Q9g997	gallus galli
604	6	2.2	622	1	DNAX_CLOPE	Q892x0	azotirizobiu	677	6	2.2	784	1	TRP1_CHICK	Q9g997	gallus galli
605	6	2.2	622	1	DNAX_CLOPE	Q892x0	azotirizobiu	678	6	2.2	784	1	TRP1_CHICK	Q9g997	gallus galli
606	6	2.2	626	1	HTRG_BACGU	Q99376	rattus norv	679	6	2.2	784	1	TRP1_CHICK	Q9g997	gallus galli
607	6	2.2	631	1	VG42_HAENI	P46208	bacillus su	680	6	2.2	787	1	TRP1_CHICK	Q9g997	gallus galli
608	6	2.2	633	1	VG42_HAENI	P46208	bacillus su	681	6	2.2	787	1	TRP1_CHICK	Q9g997	gallus galli
609	6	2.2	641	1	CGSN_PSEAE	Q50274	haemophilus	682	6	2.2	788	1	TRP1_CHICK	Q9g997	gallus galli
610	6	2.2	644	1	CGSN_PSEAE	Q50274	haemophilus	683	6	2.2	788	1	TRP1_CHICK	Q9g997	gallus galli
611	6	2.2	652	1	NU5C_PSEBK	Q87uy0	pseudomonas	684	6	2.2	790	1	TRP1_CHICK	Q9g997	gallus galli
612	6	2.2	658	1	NU5C_PSEBK	Q87uy0	pseudomonas	685	6	2.2	793	1	TRP1_CHICK	Q9g997	gallus galli
613	6	2.2	658	1	NU5C_PSEBK	Q87uy0	pseudomonas	686	6	2.2	793	1	TRP1_CHICK	Q9g997	gallus galli
614	6	2.2	662	1	MUTL1_NEIMA	Q9t62	neisseria m	687	6	2.2	795	1	TRP1_CHICK	Q9g997	gallus galli
615	6	2.2	662	1	MUTL1_NEIMA	Q9t62	neisseria m	688	6	2.2	797	1	TRP1_CHICK	Q9g997	gallus galli
616	6	2.2	667	1	SVG63_HSV1	Q00159	ictalurid h	689	6	2.2	802	1	BOB2_ACEXY	Q99k46	mus musculu
617	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	690	6	2.2	804	1	BOB2_ACEXY	Q99k46	mus musculu
618	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	691	6	2.2	805	1	BOB2_ACEXY	Q99k46	mus musculu
619	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	692	6	2.2	807	1	BOB2_ACEXY	Q99k46	mus musculu
620	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	693	6	2.2	809	1	BOB2_ACEXY	Q99k46	mus musculu
621	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	694	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
622	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	695	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
623	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	696	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
624	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	697	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
625	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	698	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
626	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	699	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
627	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	700	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
628	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	701	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
629	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	702	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
630	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	703	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
631	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	704	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
632	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	705	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
633	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	706	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
634	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	707	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
635	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	708	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
636	6	2.2	672	1	HYFB_ECOLI	Q9dbd0	mus musculu	709	6	2.2	811	1	BOB2_ACEXY	Q99k46	mus musculu
6															

691	2.2	813	1	TIRE_ECOLI	Q47281	escherichia	764	1111	1	NALS_MOUSE	Q911m5	mus musculus
692	2.2	815	1	GYRE_MYXA	Q33167	myxococcus	765	1126	1	HEM_DROME	P55162	drosophila
693	2.2	819	1	ION_CHLN	Q929f4	chlamydia p	766	1169	1	SUV3_DROME	P20193	drosophila
694	2.2	824	1	DROL_METVO	P52025	methanococc	767	1176	1	VPS8_YEAST	P39702	saccharomyc
695	2.2	826	1	SSP_PLAYO	O01443	plasmodium	768	1194	1	DPOL_VZVD	P09282	varicella-z
696	2.2	827	1	CADR_RAT	P54581	rattus norv	769	1194	1	KANK_HUMAN	O14678	homo sapien
697	2.2	829	1	IF2_HABIN	P43423	haemophilus	770	1210	1	ICEN_PSEFI	P09815	pseudomonas
698	2.2	830	1	EF2_DICDI	P15112	dictyostell	771	1238	1	DPC1_MOUSE	P54099	mus musculus
699	2.2	830	1	MCW2_SCHPO	P40377	schizosacch	772	1243	1	DPOA_ORISA	P27285	coryza sativ
700	2.2	832	1	UREA_CRYNE	O13465	cryptococcu	773	1245	1	POLS_SINDO	P03316	sindbis vir
701	2.2	833	1	CUT4_CAEEL	O17392	caenorhabdi	774	1248	1	TOFG_SULAC	O08562	sulfobolus
702	2.2	835	1	UREA_SCHPO	O00084	schizosacch	775	1257	1	EF2_METKA	O8txj4	methanopyru
703	2.2	836	1	CEG1_HUMAN	Q99490	homo sapien	776	1267	1	Y211_HUMAN	P17452	pasteurella
704	2.2	836	1	RPOC_CHLVU	P56300	chlorella v	777	1285	1	MIP1_SCHPO	O92610	homo sapien
705	2.2	837	1	UBPA_DICDI	O42401	dictyostell	778	1313	1	MRP4_HUMAN	P17452	pasteurella
706	2.2	837	1	NLG1_HUMAN	Q8n2q7	homo sapien	779	1325	1	CEG1_HUMAN	O8txj4	methanopyru
707	2.2	843	1	NLG1_MOUSE	Q99410	mus musculus	780	1360	1	CING_XENLA	P03316	sindbis vir
708	2.2	843	1	NIIG_RAT	Q62765	rattus norv	781	1377	1	NEO1_RAT	P03316	sindbis vir
709	2.2	849	1	AKA4_MOUSE	Q60662	mus musculus	782	1415	1	ICP4_HSYMG	P03316	sindbis vir
710	2.2	850	1	ADGH_USTMA	Q91495	estivnia citr	783	1418	1	SMC4_YEAST	Q91495	xenopus lae
711	2.2	850	1	ADGH_USTMA	Q91495	estivnia citr	784	1443	1	NEO1_CHICK	Q91495	xenopus lae
712	2.2	859	1	YD48_MYCTU	Q91495	estivnia citr	785	1443	1	ADP1_MYCCE	Q91495	xenopus lae
713	2.2	861	1	OPGH_PSEAE	Q11018	mycobacteri	786	1444	1	NEO1_CHICK	Q91495	xenopus lae
714	2.2	868	1	MCW2_YEAST	Q9hu46	pseudomonas	787	1451	1	EM30_ARATX	Q91495	xenopus lae
715	2.2	869	1	SVL_SYNY3	P29469	saccharomyc	788	1461	1	NEO1_HUMAN	Q91495	xenopus lae
716	2.2	876	1	KCH8_MOUSE	P73374	synchocyst	789	1469	1	BCL9_DROME	Q91495	xenopus lae
717	2.2	876	1	PRF2_YEAST	P59111	mus musculus	790	1486	1	NEO1_MOUSE	Q91495	xenopus lae
718	2.2	879	1	MANB_BOVIN	P20095	saccharomyc	791	1493	1	MUKB_ECOLI	Q91495	xenopus lae
719	2.2	879	1	MANB_CAPII	Q29444	bos taurus	792	1510	1	MUKB_ECOLI	Q91495	xenopus lae
720	2.2	879	1	PUP3_YEAST	Q95327	capra hircu	793	1516	1	Y819_PSEAE	Q91495	xenopus lae
721	2.2	880	1	SVL_CAUCR	Q07807	saccharomyc	794	1636	1	SN24_HUMAN	Q91495	xenopus lae
722	2.2	880	1	VEE1_ARCFU	Q9A5C1	caulobacter	795	1647	1	Y819_PSEAE	Q91495	xenopus lae
723	2.2	882	1	Y890_MYCTU	Q28851	archaeoglob	796	1679	1	FUR2_DROME	Q91495	xenopus lae
724	2.2	882	1	Y914_MYCBO	O10550	mycobacteri	797	1690	1	KE1A_HUMAN	Q91495	xenopus lae
725	2.2	882	1	YBAH_SCHPO	P59969	mycobacteri	798	1695	1	KE1A_MOUSE	Q91495	xenopus lae
726	2.2	885	1	IF2_SHEON	Q42308	schizosacch	799	1699	1	DROL_THR8	Q91495	xenopus lae
727	2.2	887	1	MTP_BOVIN	Q8eh15	sheareella	800	1807	1	TSC2_HUMAN	Q91495	xenopus lae
728	2.2	892	1	SVL_METKA	O57984	methanococc	801	1809	1	TSC2_RAT	Q91495	xenopus lae
729	2.2	898	1	ZOT1_XENTIA	P18751	xenopus lae	802	1816	1	KE1B_HUMAN	Q91495	xenopus lae
730	2.2	898	1	SECA_BORBU	O27497	borrelia bu	803	1816	1	KE1B_MOUSE	Q91495	xenopus lae
731	2.2	906	1	SECA_RICPR	Q92CA7	ricettsia	804	1887	1	FAS2_YEAST	Q91495	xenopus lae
732	2.2	907	1	ITTH_MOUSE	Q61702	mus musculus	805	1938	1	BGS4_SCHPO	Q91495	xenopus lae
733	2.2	907	1	NUOG_ECOL5	Q8XKX2	escherichia	806	1955	1	OTOF_HUMAN	Q91495	xenopus lae
734	2.2	907	1	NUOG_ECOL6	Q8Efi9	escherichia	807	1977	1	OTOF_MOUSE	Q91495	xenopus lae
735	2.2	907	1	NUOG_ECOLI	Q8Efi9	escherichia	808	1997	1	GLTI_YEAST	Q91495	xenopus lae
736	2.2	907	1	NUOG_SALTY	P33900	salmonella	809	2116	1	RRPL_SYNY	Q91495	xenopus lae
737	2.2	907	1	NUOG_SHIFL	Q7UC56	shigella fl	810	2144	1	POLG_HERO1	Q91495	xenopus lae
738	2.2	911	1	ITTH_HUMAN	P19827	homo sapien	811	2194	1	POLG_HERO1	Q91495	xenopus lae
739	2.2	914	1	ITTH_MESAU	P97278	mesocricetu	812	2223	1	CCAI_HUMAN	Q91495	xenopus lae
740	2.2	920	1	UBI1_HUMAN	P51784	homo sapien	813	2284	1	POLN_SFV	Q91495	xenopus lae
741	2.2	928	1	MAV4_SCHCO	P37935	schizophyll	814	2431	1	ABEC2_HUMAN	Q91495	xenopus lae
742	2.2	937	1	A2A2_RAT	P18484	rattus norv	815	2436	1	MAPB_HUMAN	Q91495	xenopus lae
743	2.2	941	1	DNAB_RHOMR	Q30477	rhodothermu	816	2468	1	MAPB_HUMAN	Q91495	xenopus lae
744	2.2	946	1	IP3L_HUMAN	P27987	homo sapien	817	2493	1	MAPB_HUMAN	Q91495	xenopus lae
745	2.2	950	1	Y136_HUMAN	Q14149	homo sapien	818	2594	1	MAPB_HUMAN	Q91495	xenopus lae
746	2.2	952	1	IF41_YEAST	P39935	saccharomyc	819	2696	1	MAPB_HUMAN	Q91495	xenopus lae
747	2.2	960	1	DLGI_DROME	P31007	drosophila	820	2842	1	MAPB_HUMAN	Q91495	xenopus lae
748	2.2	971	1	RECK_HUMAN	O95980	homo sapien	821	2843	1	MAPB_HUMAN	Q91495	xenopus lae
749	2.2	977	1	VAG9_SCHPO	O09872	schizosacch	822	2845	1	MAPB_HUMAN	Q91495	xenopus lae
750	2.2	1005	1	DBOL_VARY	P33793	variola vir	823	2869	1	MAPB_HUMAN	Q91495	xenopus lae
751	2.2	1006	1	DBOL_VACC	P20508	vaccinia vi	824	2895	1	MAPB_HUMAN	Q91495	xenopus lae
752	2.2	1006	1	DBOL_VACC	P20508	vaccinia vi	825	2895	1	MAPB_HUMAN	Q91495	xenopus lae
753	2.2	1012	1	IF2C_PRAVU	P57997	phaseolus v	826	3172	1	MAPB_HUMAN	Q91495	xenopus lae
754	2.2	1021	1	AI41_HORSE	P18907	equus caball	827	3210	1	MAPB_HUMAN	Q91495	xenopus lae
755	2.2	1029	1	END1_YEAST	P12868	saccharomyc	828	3491	1	MAPB_HUMAN	Q91495	xenopus lae
756	2.2	1045	1	CUSA_ECOL5	O8XKX2	escherichia	829	3649	1	MAPB_HUMAN	Q91495	xenopus lae
757	2.2	1047	1	CUSA_ECOL6	O8XKX2	escherichia	830	3703	1	MAPB_HUMAN	Q91495	xenopus lae
758	2.2	1047	1	CUSA_ECOLI	P38054	escherichia	831	3988	1	MAPB_HUMAN	Q91495	xenopus lae
759	2.2	1053	1	SLPM_BACBR	P06546	dacillius dr	832	4092	1	MAPB_HUMAN	Q91495	xenopus lae
760	2.2	1082	1	RPOB_EUGER	P23579	euglena gra	833	4289	1	MAPB_HUMAN	Q91495	xenopus lae
761	2.2	1093	1	AP17_HUMAN	P55198	homo sapien	834	4349	1	MAPB_HUMAN	Q91495	xenopus lae
762	2.2	1102	1	KCH8_RAT	O9tw68	rattus norv	835	4351	1	MAPB_HUMAN	Q91495	xenopus lae
763	2.2	1102	1	RPOB_AGABT	P33539	agaricus bi	836	4351	1	MAPB_HUMAN	Q91495	xenopus lae

837	6	2.2	4473	1	PLEI CRIGR	09155	criceetus	910	5	1.8	68	1	YB10	HALN1	P17104	halobacteri
838	6	2.2	4590	1	FATH HUMAN	Q14517	homo sapien	911	5	1.8	69	1	ACP RHOSH		P12784	rhodobacter
839	6	2.2	4639	1	DYHC DROME	P37276	drosoophila	912	5	1.8	69	1	GBGT BOVIN		P51554	bov taurus
840	6	2.2	4664	1	PLEI HUMAN	O15149	homo sapien	913	5	1.8	69	1	YACH RHISM		P55390	rhizobium s
841	6	2.2	4687	1	PLEI RAT	P30427	rattus norv	914	5	1.8	70	1	BRB7 MYCJM		Q9xcde	mycobacteri
842	6	2.2	5065	1	EPPL HUMAN	P58107	homo sapien	915	5	1.8	70	1	INS TORMA		P12705	torpedo mar
843	6	2.2	5120	1	PCLO CHICK	Q24496	gallus gall	916	5	1.8	70	1	R37A SULSO		Q973qz	sulfolobus
844	6	2.2	5127	1	RY44 DROME	O24498	drosoophila	917	5	1.8	70	1	RLJ1 BIFLO		Q883p6	bitfidobacte
845	6	2.2	5147	1	FAT DROME	P33450	drosoophila	918	5	1.8	70	1	YFP3 META		Q58974	methanococc
846	6	2.2	5171	1	BPEA HUMAN	O94833	homo sapien	919	5	1.8	71	1	GWPA PLIAG		Q9c3v0	planktothri
847	6	2.2	6486	1	TYCC BREPA	O30403	b tyrocidin	920	5	1.8	71	1	VGB SPV4		P11338	spiroplasma
848	6	2.2	6548	1	EPPL MOUSE	Q90W00	mus muscucu	921	5	1.8	71	1	VTB2 XENNA		P19011	xenopus lae
849	6	2.2	7073	1	RIAB CVASA	P59641	h replicase	922	5	1.8	72	1	IFI BIFLO		Q8g3z7	bitfidobacte
850	6	2.2	7389	1	BPAL MOUSE	Q91206	mus muscucu	923	5	1.8	73	1	MPF5 AMBR		P10414	ambrosia tr
851	5	1.8		1	CR41 LITCE	P56242	litorea cae	924	5	1.8	73	1	ATP9 LYCHS		P60117	lycopersico
852	5	1.8		1	CR42 LITCE	P56243	litorea cae	925	5	1.8	74	1	ATP9 MARZE		P00840	zea mays (m
853	5	1.8		1	CR43 LITCE	P56244	litorea cae	926	5	1.8	74	1	ATP9 MARZO		P26855	marcantaia
854	5	1.8		1	BRIC RANSP	P82906	rana spheno	927	5	1.8	74	1	ATP9 OENBI		P60115	oenothera b
855	5	1.8		1	FEDG AMTME	P80707	amycolatops	928	5	1.8	74	1	ATP9 PERTY		P60114	petunia hyb
856	5	1.8		1	ACP ACICA	P80916	actinobact	929	5	1.8	74	1	ATP9 SOLTU		P60114	solanum tub
857	5	1.8		1	ALR PSBFL	P33162	pseudomonas	930	5	1.8	74	1	ATP9 TOBAC		P60116	nicotiana t
858	5	1.8		1	PORA METTM	P80900	methanobact	931	5	1.8	74	1	ATP9 WHEAT		P13547	trititum ae
859	5	1.8		1	YFHA KLEOX	P21710	klebsiella	932	5	1.8	74	1	RK14 OENAM		P42340	oenothera d
860	5	1.8		1	YA79 ARCFU	O29184	archaeoglob	933	5	1.8	75	1	ATPL RHORU		P15014	rhodospirill
861	5	1.8		1	CY07 VIOOD	P58439	viola odora	934	5	1.8	75	1	HOLI BPRIT		Q38134	bacterioph
862	5	1.8		1	CECC HELVI	P83415	heliolithis v	935	5	1.8	75	1	LAFA LACCO		P24022	lactobacill
863	5	1.8		1	CECA HELVI	P83413	heliolithis v	936	5	1.8	76	1	ACP ANAVA		P20803	anabaena va
864	5	1.8		1	Y812 PASMU	Q9cm88	pasteurella	937	5	1.8	76	1	ACP CLOPE		Q8x1n7	clostridium
865	5	1.8		1	RL11 STRGB	P52661	streptomyces	938	5	1.8	76	1	ACP HAEIN		P43709	haemophilus
866	5	1.8		1	ACP PHOPR	Q9r5z3	photobacter	939	5	1.8	76	1	ACP LEUMU		P86920	leucothrix
867	5	1.8		1	CSPX KLEPN	Q48493	klebsiella	940	5	1.8	76	1	ACP OCELI		P80922	oceanospiri
868	5	1.8		1	YB8U BACSU	Q34568	bacillus su	941	5	1.8	76	1	ACP PASMU		Q9c5f5	pasteurella
869	5	1.8		1	YK82 ARCFU	O28197	archaeoglob	942	5	1.8	76	1	ACP THEIN		Q8x9w1	thermoanaer
870	5	1.8		1	R333 LACLA	Q9cd45	lactococcus	943	5	1.8	76	1	ACP VIBPA		P55337	vibrio para
871	5	1.8		1	Y5K8 SSV1	P20204	sulfolobus	944	5	1.8	76	1	TEGP HSVEA		P18554	helioobacte
872	5	1.8		1	ME23 EUPRA	P58547	euploies ra	945	5	1.8	76	1	Y58A HELPY		P57798	helioobacte
873	5	1.8		1	RL33 AERPE	P59472	aeropyrum p	946	5	1.8	77	1	ACP1 PSBEA		O54439	pseudomonas
874	5	1.8		1	RL39 SUITO	Q97111	sulfolobus	947	5	1.8	77	1	ACP AGRTS		Q8bge2	agrobacteri
875	5	1.8		1	FERA SUIME	P81843	sulfolobus	948	5	1.8	77	1	ACP CLOAB		Q971a5	clostridium
876	5	1.8		1	ATP8 POLOR	Q95912	polypetereus	949	5	1.8	77	1	ACP COMTE		P80918	comamonas t
877	5	1.8		1	COP6 STRAU	P25521	staphylococ	950	5	1.8	77	1	ACP SCOLI		P02901	escherichia
878	5	1.8		1	MAJE PSERI	P80693	pseudomonas	951	5	1.8	77	1	ACP HAEDU		Q7vkh6	haemophilus
879	5	1.8		1	RL33 ANASP	O8YGV6	anabaena sp	952	5	1.8	77	1	ACP PSBSM		B80923	pseudomonas
880	5	1.8		1	YA90 ARCFU	O29175	archaeoglob	953	5	1.8	77	1	ACP RHILE		Q9rg22	rhizobium l
881	5	1.8		1	SER4 DROME	Q9ve42	drosophila	954	5	1.8	77	1	ACP RHIME		P19372	rhizobium m
882	5	1.8		1	YAB7 ARCFU	O29178	archaeoglob	955	5	1.8	77	1	ACP SHEON		Q9edh4	shewanella
883	5	1.8		1	AKH2 LOOMI	P08379	locusta mrg	956	5	1.8	77	1	ACP SYN3		P20804	synecocyst
884	5	1.8		1	ASSY LACLO	O29173	archaeoglob	957	5	1.8	77	1	ACP VIBCH		Q9Kqhe	vibrio chol
885	5	1.8		1	YA93 ARCFU	O29173	archaeoglob	958	5	1.8	77	1	ACP VIBMA		Q9d8g3	vibrio mari
886	5	1.8		1	YML2 ARCFU	O28071	archaeoglob	959	5	1.8	77	1	ACP VIBVU		Q8z7e4	yersinia pe
887	5	1.8		1	YA95 ARCFU	O29170	archaeoglob	960	5	1.8	77	1	ACP YERPE		Q9z7e4	yersinia pe
888	5	1.8		1	CE32 HYPCU	P50722	hyphantria	961	5	1.8	77	1	IFIC ANTMA		Q94p12	antitritium
889	5	1.8		1	CE36 HYPCU	P50723	hyphantria	962	5	1.8	77	1	IFIC BOROF		Q95gm7	borago offi
890	5	1.8		1	DI07 HUMAN	Q818n7	homo sapien	963	5	1.8	77	1	IFIC CORMA		Q95gm7	corvus mas
891	5	1.8		1	RL29 NEIMA	Q91xk4	neisseria m	964	5	1.8	77	1	IFIC GAREL		Q95gm7	garrus mas
892	5	1.8		1	RPO9 MYXVL	Q9GDP6	myxoma viru	965	5	1.8	77	1	IFIC LEUPR		Q95gm2	leucophyllu
893	5	1.8		1	RPO9 SFVKA	Q9G821	shope fibro	966	5	1.8	77	1	IFIC SAMCA		Q95gm5	sembucus ca
894	5	1.8		1	YA84 ARCFU	O29181	archaeoglob	967	5	1.8	77	1	TAC1 TACTR		P14213	tachypleus
895	5	1.8		1	RL35 CHLCV	O822B3	chlamydomp	968	5	1.8	77	1	TAC2 TACTR		Q30050	archaeoglob
896	5	1.8		1	RL35 CHLCV	O84C41	chlamydia t	969	5	1.8	77	1	Y188 ARCFU		O88656	archaeoglob
897	5	1.8		1	Y314 ARCFU	O29931	archaeoglob	970	5	1.8	77	1	Y314 ARCFU		P44133	azospirillu
898	5	1.8		1	Y316 ARCFU	O29929	archaeoglob	971	5	1.8	78	1	ACP AZOBR		Q9vf55	bucella me
899	5	1.8		1	Y474 ARCFU	O29189	archaeoglob	972	5	1.8	78	1	ACP BRUME		P58449	buchnera ap
900	5	1.8		1	RL35 ARCFU	P57227	buchnera ap	973	5	1.8	78	1	ACP NEIVA		Q9j7r2	neisseria m
901	5	1.8		1	RL35 ARCFU	P49240	buchnera ap	974	5	1.8	78	1	ACP RHILU		Q884t3	rhizobium l
902	5	1.8		1	RL35 ARCFU	O895v9	buchnera ap	975	5	1.8	78	1	ACP WIGBR		P88966	xanthomonas
903	5	1.8		1	RS27 THEAC	Q9h1x2	thermoplasm	976	5	1.8	78	1	ACP XANAC		O9ezil	xanthomonas
904	5	1.8		1	RS27 THEAC	Q97Bp1	thermoplasm	977	5	1.8	78	1	ACP XANAC		P04039	saccharomyc
905	5	1.8		1	SECB THEMA	P35874	thermotoga	978	5	1.8	78	1	COX8 YEAST		P19133	sus scrofa
906	5	1.8		1	Y319 ARCFU	O29926	archaeoglob	979	5	1.8	78	1	FRIL PIG		P02937	escherichia
907	5	1.8		1	R24E HALMA	P14116	halocarcula	980	5	1.8	78	1	MULI ECOLI		P02939	erwinia amy
908	5	1.8		1	YKCD BACSU	P37480	bacillus su	981	5	1.8	78	1	MULI ECOLI		P02939	erwinia amy
909	5	1.8		1	RPO2 RALSO	Q8XG90	raistonia s	982	5	1.8	78	1	MULI ECOLI		P02939	erwinia amy

ALIGNMENTS

ID	PLPA PASHA	STANDARD;	PRT;	277 AA
----	------------	-----------	------	--------

family -

12.3%; Score 34; DB 1; Length 277

ID	METQ_HAEIN	STANDARD;	PRT;	273	AA.
----	------------	-----------	------	-----	-----

CC P31728; (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable D-methionine-binding lipoprotein metQ precursor
 DE (28 kDa outer membrane protein).
 DE MEMO OR HLPa OR H10620.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OC NCBI_TaxId=727;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Serotype B;
 RC MEDLINE=9110034; PubMed=1987077;
 RX Changyung M., Smith A.L., Moseley S.L., Kuehn M., Jenny P.;
 RT "Contribution of a 28-kDa outer membrane protein to the virulence of
 RT Haemophilus influenzae." (1991).
 RL Infect. Immun. 59:600-608(1991).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Rd / KM20 / ATCC 51907;
 RC MEDLINE=95350630; PubMed=7542800;
 RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RX Knelave A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
 RX McEwen K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RX Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RX Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RX Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,
 RX Fife L.D., Fitchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
 RX Graham C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd." (1995).
 RL Science 269:496-512(1995).
 RN [3]
 RN IDENTIFICATION BY MASS SPECTROMETRY.
 RP MEDLINE=20137488; PubMed=10675023;
 RX Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
 RA Gray C., Fountoulakis M.;
 RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
 RL Electrophoresis 21:411-429(2000).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC perase, a binding protein-dependent, ATP-driven transport
 CC system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor (Probable).
 CC -1- SIMILARITY: Belongs to the nlpa lipoprotein family.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59804; AAA24939.1; -
DR EMBL; U32744; AAC22279.1; -
DR PIR; B64082; B64082.
DR TIGR; H10620; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRfams; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DR Transprot; Amino-acid transport; Outer membrane; Lipoprotein; Signal;
DR Complete proteome; Palmitate.
DR SIGNAL 1 20 PROBABLE.
FT CHAIN 21 273 METO.
FT LIPID 21 21 N-palmitoyl cysteine (Probable).
FT LIPID 21 21 S-diacylglycerol cysteine (Probable).
FT CONFLICT 249 249 I -> V (IN REF. 1).
SQ SEQUENCE 273 AA; 29828 MW; 429838A8AC7DD7D7 CRC64;

Query Match 6.9%; Score 19; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.3e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFVYPLAGYS 125
DB 104 LNNLVIVGNTFVYPLAGYS 122

RESULT 3
ID PUPC_PASHA STANDARD; PRT; 263 AA.
AC Q08870; Q07365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane lipoprotein 3 precursor (PUP3).
GN PUPC.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype A1;
RX MEDLINE=93328110; PubMed=8335249;
RA Murphy G.L.; Whitworth L.C.;
RT "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins
RT in Pasteurella haemolytica A1."
RL Gene 129:107-111(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype A1;
RX MEDLINE=9401378; PubMed=8406866;
RA Cooney B.J.; Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
RT which are homologous to a lipoprotein gene in Haemophilus influenzae
RT type B."
RL Infect. Immun. 61:4682-4688(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC -----
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11037; AAA25540.1; -
DR EMBL; L16627; AAA25548.1; ALT_INIT.
DR PIR; JN0753; JN0753.
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR004478; YaeC.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR TIGRfams; TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DR Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 263 OUTER MEMBRANE LIPOPROTEIN 3.
FT LIPID 20 20 N-palmitoyl cysteine (Probable).
FT LIPID 20 20 S-diacylglycerol cysteine (Probable).
FT CONFLICT 103 103 T -> S (IN REF. 2).
FT CONFLICT 200 208 QNGILPTKD -> KTVANOS (IN REF. 2).
SQ SEQUENCE 263 AA; 29093 MW; 66FFD8460ED579CB CRC64;

Query Match 4.3%; Score 12; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 NLARALLILEKQ 159
DB 135 NLARALLILEKQ 146

RESULT 4
ID METQ_PASMU STANDARD; PRT; 276 AA.
AC Q9CK95;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable D-methionine-binding lipoprotein metQ precursor
DE (Outer membrane lipoprotein 1).
GN METQ OR PIPA OR PM1730.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J.; Zhang Q.; Li L.L.; Paustian M.L.; Whittam T.S.; Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: This protein is a component of a D-methionine
CC penicillase, a binding protein-dependent, ATP-driven transport
CC system (by similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006210; AA03814.1; -
DR InterPro; IPR004872; Lipoprotein_9.
DR InterPro; IPR000437; Prok_lipoprot_S.

```

DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Outer membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1 20 PROBABLE D-METHIONINE-BINDING LIPOPROTEIN
 FT CHAIN 21 276
 FT LIPID 21 21
 FT LIPID 21 21 N-palmitoyl cysteine (Probable).
 FT LIPID 21 21 S-diacylglycerol cysteine (Probable).
 SQ SEQUENCE 276 AA; 30232 MW; 7AFPD62A687D624 CRC64;

Query Match 4.3%; Score 12; DB 1; Length 276;
 Best Local Similarity 100.0%; Pred. No. 0.00052;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LLEKOGILKLD 166
 |||||
 Db 155 LLEKOGILKLD 166

RESULT 5
 METQ_ECO57 STANDARD; PRT; 271 AA.
 ID METQ_ECO57
 AC Q8X8V9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR Z0209 OR ECS0199.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Pena N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 permease, a binding protein-dependent, ATP-driven transport
 system (by similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Probable).
 CC -1- MISCELLANEOUS: The metNIQ system is also to be able to transport
 the toxic methionine analog alpha-methyl-methionine (by
 similarity).
 CC -1- SIMILARITY: Belongs to the nlpA lipoprotein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE005195; AGS4499.1;
 DR EMBL: AP002550; BAB33622.1;
 DR PIR: G85504; G85504.
 DR PIR: G90653; G90653.
 DR InterPro: IPR004872; Lipoprotein 9.
 DR InterPro: IPR00437; ProX_Lipoprot_S.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 KM Complete proteome; Palmitate.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
 FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 SQ SEQUENCE 271 AA; 29416 MW; 39D7DA1B5DCD2B8B CRC64;

Query Match 4.0%; Score 11; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.0049;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKSPYVN 234
 |||||
 Db 219 FVEDKSPYVN 229

RESULT 6
 METQ_SALTI STANDARD; PRT; 271 AA.
 ID METQ_SALTI
 AC Q8Z892;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR STY0272 OR T0248.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Iiou S.-R., Plunkett G., III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 and CT18";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 permease, a binding protein-dependent, ATP-driven transport
 system (by similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Probable).
 CC -1- MISCELLANEOUS: The metNIQ system is also to be able to transport

CC the toxic methionine analog alpha-methyl-methionine (By
CC similarity).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL627266; CAD08705.1; -
CC EMBL: AB016834; AAC67977.1; -
CC InterPro: IPR004872; Lipoprotein_9.
CC InterPro: IPR000437; Prok_lipoprot_5.
CC InterPro: IPR004478; Yaec.
CC Pfam: PF03180; Lipoprotein_9; 1.
CC TIGRfam: TIGR00363; TIGR00363; 1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Transport: Amino-acid transport; Membrane; Lipoprotein; Signal;
CC Complete proteome; Palmitate.
CC SIGNAL 1 22
CC CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
CC LIPID 23 23 N-palmitoyl cysteine (Potential).
CC LIPID 23 23 S-diacylglycerol cysteine (Potential).
CC SEQUENCE 271 AA; 29465 MW; 49B62C4CF96D9613 CRC64;
SQ
Query Match 4.0%; Score 11; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 224 FVEDKDSPLYN 234
DB 219 FVEDKDSPLYN 229
RESULT 7
METQ_SALTY STANDARD; PRT; 271 AA.
ID METQ_SALTY STANDARD; PRT; 271 AA.
AC 082KNT;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE D-methionine-binding lipoprotein metq precursor.
GN METQ OR STM0245.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RX MEDLINE=21534948; PubMed=1677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
CC -1- FUNCTION: This protein is a component of a D-methionine
CC permease, a binding protein-dependent, ATP-driven transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -1- MISCELLANEOUS: The metnIQ system is also to be able to transport
CC the toxic methionine analog alpha-methyl-methionine (By
CC similarity).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB008706; AAL19208.1; -
CC STyGene: SG2727; metQ.
CC InterPro: IPR004872; Lipoprotein_9.
CC InterPro: IPR000437; Prok_lipoprot_5.
CC InterPro: IPR004478; Yaec.
CC Pfam: PF03180; Lipoprotein_9; 1.
CC TIGRfam: TIGR00363; TIGR00363; 1.
CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
CC Transport: Amino-acid transport; Membrane; Lipoprotein; Signal;
CC Complete proteome; Palmitate.
CC SIGNAL 1 22
CC CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
CC LIPID 23 23 N-palmitoyl cysteine (Potential).
CC LIPID 23 23 S-diacylglycerol cysteine (Potential).
CC SEQUENCE 271 AA; 29435 MW; 3FED41E14FBD707 CRC64;
SQ
Query Match 4.0%; Score 11; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 224 FVEDKDSPLYN 234
DB 219 FVEDKDSPLYN 229
RESULT 8
RPA9_YEAST STANDARD; PRT; 125 AA.
ID RPA9_YEAST STANDARD; PRT; 125 AA.
AC P32529;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DNA-directed RNA polymerase I 13.7 kDa polypeptide (EC 2.7.7.6)
DE (A12.2).
GN RPA12 OR RRM4 OR YTR063M OR J1747.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RX MEDLINE=93109294; PubMed=8417319;
RA Nogai Y., Yano R., Dodd J., Carles C., Nomura M.;
RT "Gene RRM4 in Saccharomyces cerevisiae encodes the A12.2 subunit of
RT RNA polymerase I and is essential only at high temperatures.";
RL Mol. Cell. Biol. 13:114-122 (1993).
RN [1]
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RP MEDLINE=93109294; PubMed=8417319;
RA Nogai Y., Yano R., Dodd J., Carles C., Nomura M.;
RC STRAIN=S288C;
RX MEDLINE=96437976; PubMed=8840504;
RA Huang M.-E., Manns V., Chant J.-C., Galibert F.;
RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT reading frames and a gene cluster with a counterpart on chromosome
RT XI.";
RL Yeast 12:869-875 (1996).
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: RNA polymerase I consists of 14 different subunits.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase

-1- SIMILARITY: Belongs to the archaeal rpm / eukaryotic RPA12/RP29/RPCL RNA polymerase family.
 -1- SIMILARITY: Contains 1 ribbon-type zinc finger.

 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; L00708; AAA3492.1; -
 CC EMBL; L35564; AAB59319.1; -
 CC EMBL; L49563; CAA89591.1; -
 CC EMBL; L47993; AAB39289.1; -
 CC PIR; A48107; A48107.
 CC HSP; Q56254; IQP.
 CC Germonline; 141896; -
 CC SGD; S0003824; RPA12.
 CC InterPro; IPR001529; RNA POL.
 CC InterPro; IPR001222; TFIS.
 CC Pfam; PF02150; RNA_POL_M_15KD; 1.
 CC Pfam; PF01096; TFIS; 1.
 CC SMART; SM00661; RPO19; 1.
 CC SMART; SM00440; ZNF_C2C2; 1.
 CC PROSITE; PS00466; TFIS; 1.
 CC PROSITE; PS01030; RNA_POL_M_15KD; 1.
 CC KX Transferase; DNA-directed RNA polymerase; Transcription;
 CC Nucleic protein; zinc-finger.
 CC ZN_FING 10 33 ZN-RIBSON (POTENTIAL).
 CC ZN_FING 86 117 ZN-RIBSON (POTENTIAL).
 CC SQ SEQUENCE 125 AA; 13660 MW; D79372070819687C CRC64;

Query Match 3.3%; Score 9; DB 1; Length 125;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 NELKDGATI 140
 DB 74 NELKDGATI 82

 CC RESULT 9
 CC METQ_VIBCH STANDARD; PRT; 269 AA.
 CC ID METQ_VIBCH
 CC AC O9KJ77;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Probable D-methionine-binding lipoprotein meto precursor.
 CC GN METQ OR VC0905.
 CC OS Vibrio cholerae.
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 CC OC Vibrionaceae; Vibrio.
 CC OX NCBI_Taxid=666;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=EI TOR N16961 / Serotype O1;
 CC RA MEDLINE=20406833; PubMed=10952301;
 CC RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 CC Dooson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 CC Gill S.R., Nelson K.E., Read T.D., Tettein H., Richards D.,
 CC Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Delo I., Sellers P.,
 CC McDonald L., Ureback T., Fleischmann R.D., Nieman W.C., White O.,
 CC Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 CC Fraser C.W.;
 CC RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 CC RT cholerae.";
 CC RL Nature 406:477-483 (2000).
 CC CC -1- FUNCTION: This protein is a component of a D-methionine
 CC pemase: a binding protein-dependent, ATP-driven transport
 CC system (by similarity).

-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.

 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AE004174; AAF94067.1; ALT_INIT.
 CC TIGR; VC0905; -
 CC InterPro; IPR004872; Lipoprotein_9.
 CC InterPro; IPR000437; Prok_Lipoprot_S.
 CC InterPro; IPR004478; YaeC.
 CC Pfam; PF01180; Lipoprotein_9; 1.
 CC TIGRFAMs; TIGR00363; TIGR00363; 1.
 CC PROSITE; PS00013; PROKR_LIPOPROTEIN; 1.
 CC KX Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 CC Complete proteome; Palmitate.
 CC FT SIGNAL 1 22 POTENTIAL.
 CC FT CHAIN 23 269 PROBABLE D-METHIONINE-BINDING LIPOPROTEIN
 CC FT METQ.
 CC FT LIPID 23 N-palmitoyl cysteine (Potential).
 CC FT LIPID 23 S-diacylglycerol cysteine (Potential).
 CC SQ SEQUENCE 269 AA; 29067 MW; 39E2570E33F184D6 CRC64;

Query Match 3.3%; Score 9; DB 1; Length 269;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 IVGNTFYYP 120
 DB 105 IVGNTFYYP 113

 CC RESULT 10
 CC MCB_HUMAN STANDARD; PRT; 391 AA.
 CC ID MCB_HUMAN
 CC AC P04220;
 CC DT 20-MAR-1987 (Rel. 04, Created)
 CC DT 20-MAR-1987 (Rel. 04, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Ig mu heavy chain disease protein (BOR).
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC OX NCBI_Taxid=9606;
 CC RN [1]
 CC RP SEQUENCE.
 CC RX MEDLINE=84184186; PubMed=6425189;
 CC RA Barnikol-Watanabe S., Milnesco E., Milnesco C., Barnikol H.U.,
 CC Hirschman N.;
 CC RA "The primary structure of mu-chain-disease protein BOR. Peculiar
 CC RT amino-acid sequence of the N-terminal 42 positions."
 CC HL Hoppe-Seyler's Z. Physiol. Chem. 365:105-118 (1984).
 CC CC -1- MISCELLANEOUS: This protein has no V region homology or CH1
 CC region.
 CC PIR; A02163; MHHTBT.
 CC HSP; F01857; 1FCT1.
 CC GO; GO:0005624; C:membrane fraction; NAS.
 CC GO; GO:0003823; P:antigen binding; NAS.
 CC GO; GO:0006955; P:immune response; NAS.
 CC DR InterPro; IPR007110; IG-like.
 CC DR InterPro; IPR003597; IG cl.
 CC DR InterPro; IPR003006; IG_MHC.
 CC DR Pfam; PF00477; Ig; 3.
 CC DR SMART; SM00407; IGcl; 2.
 CC DR PROSITE; PS00835; IG_LIKE; 3.
 CC DR PROSITE; PS00290; IG_MHC; 3.
 CC KM Immunoglobulin domain; Immunoglobulin C region.

FT DOMAIN 1 42 PRE-C-PART (NO V REGION HOMOMLOGY).
 FT DOMAIN 43 155 CH2.
 FT DOMAIN 156 261 CH3.
 FT DOMAIN 262 391 CH4.
 SQ SEQUENCE 391 AA; 43057 MW; 9100843AFC021A CRC64;

Query Match 3.3%; Score 9; DB 1; Length 391;
 Best Local Similarity 100.0%; Pred. No. 0.64;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 TDEVEARAK 265
 DB 99 TDEVEARAK 107

RESULT 11
 YD97_CAMD 1
 ID YD97_CAMD 1 STANDARD; PRT; 74 AA.
 AC 09PMEO;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein Cj1397.
 GN Cj1397.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jorgensen K., Kariyasek A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RT Nature 403:665-668(2000).
 RL Nature 403:665-668(2000).
 CC -1-SIMILARITY: TO H. PYLORI HP0585.1/HP0533.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AJ139078; CAB73821.1; -;
 DR EMBL: AJ139078; CAB73821.1; -;
 DR PIR: E81284; E81284.
 DR InterPro: IPR007167; FeoA.
 DR Pfam: PF04023; FeoA; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 74 AA; 8363 MW; E81BC8AB369DDC4 CRC64;

Query Match 2.9%; Score 8; DB 1; Length 74;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 TUNELKDG 137
 DB 2 TUNELKDG 9

RESULT 12
 METQ_ECOLI 1
 ID METQ_ECOLI 1 STANDARD; PRT; 271 AA.
 AC P28635;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1993 (Rel. 26, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE D-methionine-binding lipoprotein metQ precursor.
 GN METQ OR B0197.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Miyamoto K.;
 RL Submitted (Apr-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RL "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Samei G., Mizobuchi K.;
 RL "Systematic sequencing of the Escherichia coli genome: analysis of the
 RL 4.0 - 6.0 min (189,987 - 281,416bp) region.";
 RL Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurd O.,
 RA Laethkari D., Lew H., Lin D., Namath A., Oetner P., Roberts D.,
 RA Davis R.W.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 127-271 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93094132; PubMed=1459951;
 RA Gervais F.G., Drepeau G.R.;
 RL "Identification, cloning, and characterization of rcsF, a new
 RL regulator gene for exopolysaccharide synthesis that suppresses the
 RL division mutation fts284 in Escherichia coli K-12.";
 RL J. Bacteriol. 174:8016-8022 (1992).
 RN [6]
 RP FUNCTION.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=22159950; PubMed=12169620;
 RA Gal J., Szvechnik A., Schneid R., Kalman M.;
 RL "The metQ D-methionine transporter locus of Escherichia coli is an ABC
 RL transporter gene cluster.";
 RL J. Bacteriol. 184:4930-4932 (2002).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- MISCELLANEOUS: The metQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine.
 CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
 CC -1- CAUTION: REF.5 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITION 217
 CC ONWARD DUE TO A FRAMESHIFT.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AL133215; -; NOT ANNOTATED CDS.
 CC EMBL; AB046839; BAB13445.1; ALT_INIT.
 CC Genew; HGNCL10735; SEMA4G.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003599; Ig.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; Ig; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG_Like; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 CC Neurogenesis; Developmental protein; Glycoprotein.
 CC SIGNA; 1
 CC CHAIN; 17
 CC FT; 18 838
 CC FT; 18 675
 CC FT; 18 696
 CC FT; 697 838
 CC FT; 56 489
 CC FT; 507 558
 CC FT; 567 649
 CC FT; 565 568
 CC FT; 763 774
 CC FT; 584 632
 CC FT; 55 55
 CC FT; 111 111
 CC FT; 126 126
 CC FT; 388 388
 CC FT; 542 542
 CC FT; 598 598
 CC FT; 543 543
 CC FT; 543 543
 CC SQ SEQUENCE 838 AA; 91496 MW; 9B281AE8681F245 CRC64;

Query Match 2.9%; Score 8; DB 1; Length 838;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 PSMILARAL 153
 DB 587 PSMILARAL 594

RESULT 15
 ID LBN MOUSE STANDARD; PRT; 1220 AA.
 AC ORKIG2; Q8BRF3;
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lamin.
 GN EVC2 OR LBN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Limb bud;
 RX MEDLINE=22155879; Pubmed=12136126;
 RA Takea H., Takami M., Oguni T., Tsuji T., Yoneda K., Sato H.,
 RA Ihara N., Itoh T., Kato S.R., Mishina Y., Womack J.E., Moritomo Y.,
 RA Sugimoto Y., Kunieda T.,
 RA "Positional cloning of the gene lamin responsible for bovine
 RT chondrodysplastic dwarfism";

RL Proc. Natl. Acad. Sci. U.S.A. 99:10549-10554 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RX MEDLINE=22388257; Pubmed=12477932;
 RA Strussberg R.L., Feinsgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Holys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman D.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RP SEQUENCE OF 21-744 FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=22354683; Pubmed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nixaido I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach U.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Datta E., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G.,
 RA Petrovsky N., Pillai R., Pontius U.V., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wainwright C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carrinci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Maki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs";
 Nature 420:563-573 (2002).
 RL -1- FUNCTION: Plays a critical role in bone formation and skeletal
 CC development. May be involved in early embryonic morphogenesis.
 CC -1- TISSUE SPECIFICITY: Expressed in long and cranial bones, kidney
 CC and heart. Strongly expressed in proliferating chondrocytes,
 CC osteoblasts and osteoclasts.
 CC -1- DEVELOPMENTAL STAGE: Found in the embryo at day E7, E11, E15, and
 CC E17. At the limb bud formation stage E11, it is expressed in fore-
 CC and hindlimb buds, branchial arches, and facial primordia.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB083066; BAC06589.1; -
CC DR EMBL; BC037473; AAC37473.1; -
CC DR EMBL; AK044977; BAC32167.1; ALT_INIT.
CC DR MGD; MGI:191575; Evc2.
CC KM Transmembrane; Colled coil.
CC FT DOMAIN 211 231 POTENTIAL.
CC FT DOMAIN 355 404 COILED COIL (POTENTIAL).
CC FT DOMAIN 563 644 COILED COIL (POTENTIAL).
CC FT DOMAIN 854 875 COILED COIL (POTENTIAL).
CC FT DOMAIN 920 1005 COILED COIL (POTENTIAL).
CC SQ SEQUENCE 1220 AA; 137638 MW; E6767171A986BD CRC64;

Query Match 2.9%; Score 8; DB 1; Length 1220;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 SAVSKGEL 87
Db 457 SAVSKGEL 464

RESULT 16
ID HFB_XANAC STANDARD; PRT; 103 AA.
AC O8P8F6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integration host factor beta-subunit (IHF-beta).
GN IHF OR HIMD OR XAC2297.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Canargo L.E.A.,
RA Canarotte G., Canavan F., Cardozo J., Chamberggo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.S., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC translational control (By similarity).
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain (By similarity).
CC -!- SIMILARITY: Belongs to the bacterial histone-like protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; AE011966; AAM37150.1; -
DR HAVAP; MF_00381; -; 1.
DR InterPro; IPR000119; Bac DNABind.
DR InterPro; IPR005685; IHF_beta.
DR Pfam; PF00216; Bac DNA binding; 1.
DR PRINTS; PRO1727; DNABINDINGHU.
DR ProDom; PD000945; Bac DNABind; 1.
DR TIGRfam; TIGR00388; hfp; 1.
DR PROSITE; PS00045; HISTONE_LIKE; 1.
KM DNA-binding; Transcription regulation; DNA recombination;
KM translation regulation; Complete proteome.
SQ SEQUENCE 103 AA; 11373 MW; 8D5F59687D3FDB6 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 DDVDLAV 205
Db 20 DDVDLAV 26

RESULT 17
ID HFB_XANCP STANDARD; PRT; 103 AA.
AC O8P8F6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integration host factor beta-subunit (IHF-beta).
GN IHF OR HIMD OR XCC2193.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A. / NCPPB 528;
RC STRAIN=ATCC 33913;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Canargo L.E.A.,
RA Canarotte G., Canavan F., Cardozo J., Chamberggo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.S., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC translational control (By similarity).
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain (By similarity).
CC -!- SIMILARITY: Belongs to the bacterial histone-like protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012326; AAM41473.1; -

```

```

DR HAMAP: MF_00381; -; 1.
DR InterPro: IPR000119; Bac_DNAbind.
DR Pfam: PF00216; Bac_DNA_binding; 1.
DR PRINTS: PR01727; DNABINDINGH.
DR PRODOM: PD000945; Bac_DNAbind; 1.
DR TIGRFAMs: TIGR00988; hlp; 1.
DR PROSITE: PS00045; HISTONE LIKE; 1.
DR DNA-binding; Transcription regulation; DNA recombination;
KM Translation regulation; Complete proteome.
SQ SEQUENCE 103 AA; 11403 MW; 8D5F45186237ADB6 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 DDVDLAV 205
DB 20 DDVDLAV 26

RESULT 18
ID IHFB_XYLFA STANDARD; PRT; 104 AA.
AC QSPAC8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integration host factor beta-subunit (IHFB-beta).
GN IHFB OR HIMD OR XP2437.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=945C; PubMed=10910347;
RX MEDLINE=20365717;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares A.J.G., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Breno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.R., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facchini A.P., Ferreira A.J.S., Ferreira V.C.A., Fierro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Kieger J.E., Kuramae E.E., Laigret F., Lambis M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.X.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Mizoca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pequeiro J.B.,
RA Queiroz R.E., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silveira M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubaki M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zechin J., Zetser J.C.,
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
CC -1- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC translational control (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -1- SIMILARITY: Belongs to the bacterial histone-like protein family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004052; AAP85236.1; AUT_INIT.
DR HSSP: P36206; 1B82.
DR HAMAP: MF_00381; -; 1.
DR InterPro: IPR000119; Bac_DNAbind.
DR Pfam: PF00216; Bac_DNA_binding; 1.
DR PRINTS: PR01727; DNABINDINGH.
DR PRODOM: PD000945; Bac_DNAbind; 1.
DR SMART: SM00411; BHL; 1.
DR TIGRFAMs: TIGR00988; hlp; 1.
DR PROSITE: PS00045; HISTONE LIKE; 1.
DR DNA-binding; Transcription regulation; DNA recombination;
KM Translation regulation; Complete proteome.
SQ SEQUENCE 104 AA; 11463 MW; 779514E91FA731CC CRC64;

Query Match 2.5%; Score 7; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 DDVDLAV 205
DB 20 DDVDLAV 26

RESULT 19
ID IHFB_XYLFT STANDARD; PRT; 104 AA.
AC Q87B8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integration host factor beta-subunit (IHFB-beta).
GN IHFB OR HIMD OR PD1455.
OS Xylella fastidiosa (strain Temecual / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita W.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Petro E.S., Harakava R., Kuramae E.E.,
RA Martino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.V.,
RA Bata G.S., Bianco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsubaki M.H., Yanai G.M., Zatz M.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RL J. Bacteriol. 185:1018-1026(2003).
CC -1- FUNCTION: This protein is one of the two subunits of integration
CC host factor, a specific DNA-binding protein that functions in
CC genetic recombination as well as in transcriptional and
CC translational control (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain (By similarity).
CC -1- SIMILARITY: Belongs to the bacterial histone-like protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AEO12558; AAO29299.1; -
 CC DR HAVAP; MF 00381; -; 1.
 CC DR InterPro; IPR000119; Bac DNABind.
 CC DR Pfam; PF00216; Bac DNA binding; 1.
 CC DR PRINTS; PRO1727; DNABINDINGHU.
 CC DR PROSITE; PS00045; HISTONE_LIKE; 1.
 CC KM DNA-binding; Transcription regulation; DNA recombination;
 CC KM translation regulation; Complete proteome.
 CC SQ SEQUENCE 104 AA; 11377 MW; 999677891CC451C9 CRC64;
 CC
 CC Query Match 2.5%; Score 7; DB 1; Length 104;
 CC Best Local Similarity 100.0%; Pred. No. 19;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 199 DDVDLAV 205
 CC Db 20 DDVDLAV 26
 CC
 CC RESULT 20
 CC PSP1_PSEIC STANDARD; PRT; 141 AA.
 CC ID PSP1_PSEIC
 CC AC 061704;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Plasmacyte-spreading peptide precursor.
 CC GN PSP1.
 CC OS Pseudoplusia includens (Soybean looper).
 CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
 CC OC Noctuidae; Plusiinae; Pseudoplusia.
 CC CX NCBI_TaxID=76492;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP MEDLINE=98440825; PubMed=9753657;
 CC RA Clark K.D., Witherell A., Strand M.R.;
 CC RT "Plasmacyte spreading peptide is encoded by an mRNA differentially
 CC RT expressed in tissues of the moth Pseudoplusia includens.";
 CC RL Biochem. Biophys. Res. Commun. 250:479-485(1998).
 CC [2]
 CC RN SEQUENCE OF 119-141.
 CC RP TISSUE=Hemolymph;
 CC RC MEDLINE=97435321; PubMed=9287360;
 CC RA Clark K.D., Pech L.L., Strand M.R.;
 CC RT "Isolation and identification of a plasmacyte-spreading peptide
 CC RT from the hemolymph of the lepidopteran insect Pseudoplusia
 CC RT includens.";
 CC RL J. Biol. Chem. 272:23440-23447(1997).
 CC [3]
 CC RN STRUCTURE BY NMR OF 119-141.
 CC RP TISSUE=Hemolymph;
 CC RC MEDLINE=99143099; PubMed=9988679;
 CC RA Volkman B.F., Anderson M.E., Clark K.D., Hayakawa Y., Strand M.R.,
 CC RA Markley J.L.;
 CC RT "Structure of the insect cytokine peptide plasmacyte-spreading
 CC RT peptide 1 from Pseudoplusia includens.";
 CC RL J. Biol. Chem. 274:4493-4496(1999).
 CC -1- FUNCTION: Mediates the spreading of plasmacytes to foreign
 CC surfaces. Plasmacytes are a class of hemocytes involved in insect
 CC cellular immunity.
 CC -1- SIMILARITY: Belongs to the GBP / PSP1 / paralytic peptide family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF062489; AAC16546.1; -
 CC DR PIR; J50359; J50359.
 CC DR PDB; 1BYV; 02-DEC-98.
 CC DR PDB; 1BSN; 13-JAN-99.
 CC DR InterPro; IPR003463; GBP_PSP.
 CC DR Pfam; PF02425; GBP_PSP; 1.
 CC DR Prodom; PD06507; GBP_PSP; 1.
 CC KM Hemolymph; Signal; 3D-structure.
 CC FT SIGNAL 1 22
 CC FT PROPEP 23 118
 CC FT PEPTIDE 119 141
 CC FT DISULFID 125 137
 CC FT STRAND 130 131
 CC FT STRAND 137 138
 CC SQ SEQUENCE 141 AA; 14953 MW; 929B4BD55CP21A4 CRC64;
 CC
 CC Query Match 2.5%; Score 7; DB 1; Length 141;
 CC Best Local Similarity 100.0%; Pred. No. 25;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 223 VFVEDKD 229
 CC Db 65 VFVEDKD 71
 CC
 CC RESULT 21
 CC COAT_ACLSIA STANDARD; PRT; 193 AA.
 CC ID COAT_ACLSIA
 CC AC P54890;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 01-OCT-1996 (Rel. 34, Last annotation update)
 CC DE Coat protein.
 CC OS Apple chlorotic leaf spot virus (isolate apple) (ACLSV).
 CC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
 CC OC Trichovirus.
 CC CX NCBI_TaxID=73472;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP MEDLINE=93389448; PubMed=8376968;
 CC RA Sato K., Yoshikawa N., Takahashi T.;
 CC RT "Complete nucleotide sequence of the genome of an apple isolate of
 CC RT apple chlorotic leaf spot virus.";
 CC RL J. Gen. Virol. 74:1927-1931(1993).
 CC [2]
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; D14996; BAA03643.1; -
 CC DR PIR; JQ2185; JQ2185.
 CC DR InterPro; IPR008879; Tricho_coat.
 CC DR Pfam; PF05892; Tricho_coat; 1.
 CC KM Coat protein.
 CC SQ SEQUENCE 193 AA; 21395 MW; C1E030BR33A7COA CRC64;
 CC
 CC Query Match 2.5%; Score 7; DB 1; Length 193;
 CC Best Local Similarity 100.0%; Pred. No. 33;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 168 TNLPEST 174
 CC Db 126 TNLPEST 132

```

RESULT 22
MRKB_KLEPN STANDARD; PRT; 233 AA.
AC P21646;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Chaperone protein mrkb precursor.
MRKB.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxId=573;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=1A565;
RX MEDLINE=91100388; PubMed=1670938;
RA Allen B.L., Gerlach G.-F., Clegg S.;
RT "Nucleotide sequence and functions of mrk determinants necessary for
RT expression of type 3 fimbriae in Klebsiella pneumoniae.";
RL J. Bacteriol. 173:916-920(1991)
CC -1- FUNCTION: MEDIANES ASSEMBLY OF PILI BY FORMING SOLUBLE MULTIMERIC
CC COMPLEXES WITH PILI SUBUNITS AS AN INTERMEDIATE STEP IN THE
CC ASSEMBLY PROCESS. THIS PROTEIN IS INVOLVED IN TYPE 3 PILI
CC ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: Belongs to the periplasmic pilus chaperone family.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M55912; AAA25094.1; -.
CC PIR; C39142; C39142.
CC HSSP; P31697; 1BF8.
CC InterPro: IPR008962; PapD-like.
CC InterPro: IPR001829; Pili chaperone.
CC Pfam; PF00345; Pili_assembly; 1.
CC PRINTS; PR00969; CHAPERONPILI.
CC ProDom; PD001447; Pili chaperone; 1.
CC PROSITE; PS00635; Pili_CHAPERONE; 1.
CC Chaperone; Fimbria; Periplasmic; Signal; Immunoglobulin domain.
CC SIGNAL 1 18 POTENTIAL.
CC CHAIN 19 233 CHAPERONE PROTEIN MRKB.
CC SEQUENCE 233 AA; 25143 MW; F71EAA950AF2317C CRC64;
SQ
Query Match 2.5%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 SAVSKG3 86
DB 206 SAVSKG3 212

```

```

OX NCBI_TaxId=2097;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischman R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.D.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Boett K.F., Hu P.-C., Luster T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC -1- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the guanylate kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39691; AAC71325.1; ALT_INIT.
CC HSSP; P15454; 1GXY.
CC TIGR; MG107; -.
CC
CC HAMAP; MF_00328; -. 1.
CC InterPro: IPR008144; Guanylate_kin.
CC InterPro: IPR008145; Guanylt/Ca.
CC Pfam; PF00625; Guanylate_kin; 1.
CC SMART; SMC0072; GuKc; 1.
CC PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
CC Transferrase; Kinase; ATP-binding; Complete proteome.
CC NP_BIND 63 70
CC SEQUENCE 240 AA; 27520 MW; 0664E11FE90935D CRC64;
SQ
Query Match 2.5%; Score 7; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 129 KTLNELK 135
DB 226 KTLNELK 232

```

```

RESULT 23
KGDA_MYCGE STANDARD; PRT; 240 AA.
AC P47353;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN CMK OR MG107.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

```

```

RESULT 24
CLCR_RAT STANDARD; PRT; 268 AA.
AC P55091; Q63188;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcitriol precursor (EC 3.4.21.2) (Chymotrypsin C) (Serum calcium-
DE decreasing factor).
GN CTIRC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pancreas;
RX MEDLINE=96107178; PubMed=8530454;
RA Tomomura A., Tomomura M., Fukushima T., Akiyama M., Kubota N.,
RA Kumaki K., Nishii Y., Noikura T., Saheki T.;
RT "Molecular cloning and expression of serum calcium-decreasing factor
RT (calcitriol).";
RL J. Biol. Chem. 270:30315-30321(1995).
RN [2]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=92165057; PubMed=1537555;
RA Kang U., Wiegand U., Mueller-Hill B.;
RT "Identification of cDNAs encoding two novel rat pancreatic serine
RT processes.";
RL Gene 110:181-187(1992).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98207038; PubMed=9538241;
RA Yoshino-Yasuda I., Kobayashi K., Akiyama M., Itoh H., Tomomura A.,
RA Sabeke T.;
RT "Caldecrin is a novel-type serine protease expressed in pancreas, but
RT its homologue, elastase IV, is an artifact during cloning derived
RT from caldecrin gene.";
RL J. Biochem. 123:546-554(1998).
CC -1- FUNCTION: Has chymotrypsin-type protease activity and hypocalcemic
CC activity.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Tyr|-Xaa,
CC Phe|-Xaa, Met|-Xaa, Trp|-Xaa, Gln|-Xaa, Asn|-Xaa.
CC -1- TISSUE SPECIFICITY: Pancreas.
CC -1- SIMILARITY: Belongs to peptidase family S1. Elastase subfamily.
CC -1- CAUTION: Was originally thought to be elastase IV.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S80379; AAB35830.1; -;
DR EMBL: X59014; CAA41753.1; -;
DR PIR: J01473; J01473.
DR HSP: P00766; ICHG.
DR MEROPS: S01.157; -;
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; tryp_SPC; 1.
DR PROSITE: PS00240; TRYPsin_D0M; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydrolyse: Serine protease; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 29
FT CHAIN 30 268
FT ACT_SITE 123 123
FT ACT_SITE 126 216
FT DISULFID 17 141
FT DISULFID 59 75
FT DISULFID 155 222
FT DISULFID 186 202
FT DISULFID 212 243
FT CARBOHYD 25 25
FT CARBOHYD 90 90
FT CONFLICT 42 42
FT CONFLICT 96 120
SQ SEQUENCE 268 AA; 29374 MW; 33B67AF34D0F8583 CRC64;
TSMSGTDSSCGT (IN REF. 2).
TSMSGVAEDTVYHEKMLFLWN -> AEACTLRMTPS
Query Match 2.5%; Score 7; DB 1; Length 268;
Best Local Similarity 100.0%;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 62 KYNLTVE 68
DB 88 KYNLTVE 94

RESULT 25
ID_SAPF_ECOLI STANDARD; PRT; 268 AA.
AC P36637;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Peptide transport system ATP-binding protein sapf.
GN SAPF OR B1290 OR C1767 OR SF1295 OR S1377.
OS Escherichia coli,
OS Escherichia coli O6, and
OS Shigella flexneri.
OC Bacteria, Proteobacteria, Gammaproteobacteria, Enterobacteriales;
OC Enterobacteriaceae, Escherichia.
OX NCBI_TaxID=562, 217992, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RA Bergler H., Ebeling A., Fuchsbichler S., Hogenauer G., Turnowsky F.;
RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RA Epstein W., Neelker B., Stumpe S., Tewes R., Schmid R., Bakker E.P.;
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Iseno K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Nakade S., Nakamura Y., Nishino Y.,
RA Motomura K., Nakade S., Nakamura Y., Nishino Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubraman S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horuchi T.;
RL "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Itoh S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.G.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RL "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).

RT [7]
RP SEQUENCE FROM N.A.
RC SPECIES=flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=2550274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T".
RL Infect. Immun. 71:2775-2786(2003).
CC -1- FUNCTION: INVOLVED IN A PEPTIDE INTAKE TRANSPORT SYSTEM THAT
CC PLAYS A ROLE IN THE RESISTANCE TO ANTIMICROBIAL PEPTIDES.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08190; AAA17671.1; -
DR EMBL; X97282; CAA65941.1; -
DR EMBL; AE000227; AAC74372.1; -
DR EMBL; D90766; BAA14843.1; -
DR EMBL; D90767; BAA14851.1; -
DR EMBL; AE016760; AAN80233.1; -
DR EMBL; AE015156; AAN42506.1; -
DR EMBL; AE016982; AAP16789.1; -
DR PIR; E64877; E64877.
DR EcoGene; EG12305; sapF.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW Peptide transport; Transport; Inner membrane; ATP-binding;
KW Complete proteome.
FT NP_BIND 47
SQ SEQUENCE 268 AA; 30570 MW; F3B98BC9DC6CBFA CRC64;
FT N_BIND 47
QY Query Match 2.5%; Score 7; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 149 LARALIL 155
161 LARALIL 167

RESULT 26
SAPF_SALTY
ID SAPF_SALTY STANDARD; PRT; 268 AA.
AC P36638;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide transport system ATP-binding protein sapF.
GN SAPF OR STM1696.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14028S;
RX MEDLINE=94038887; PubMed=8223423;
RA

RT "Molecular genetic analysis of a locus required for resistance to
RT antimicrobial peptides in Salmonella typhimurium.";
RL EMBL J. 12:4053-4062(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Kulkarny E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: INVOLVED IN A PEPTIDE INTAKE TRANSPORT SYSTEM THAT
CC PLAYS A ROLE IN THE RESISTANCE TO ANTIMICROBIAL PEPTIDES.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74212; CAA52268.1; -
DR EMBL; AE008775; AAL20613.1; -
DR PIR; G39589; S39589.
DR StyGene; SG10381; sapF.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW Peptide transport; Transport; Inner membrane; ATP-binding;
KW Complete proteome.
FT NP_BIND 47
SQ SEQUENCE 268 AA; 30671 MW; 339B649A75252504 CRC64;
FT N_BIND 47
QY Query Match 2.5%; Score 7; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 149 LARALIL 155
161 LARALIL 167

RESULT 27
SAPF_HABIN
ID SAPF_HABIN STANDARD; PRT; 269 AA.
AC P45289;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide transport system ATP-binding protein sapF.
GN SAPF OR H11643.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleisemann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.W.,

RA Scott J.D., Shiley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
 RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.,
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.",
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: INVOLVED IN A PEPTIDE INTAKE TRANSPORT SYSTEM THAT
 CC PLAYS A ROLE IN THE RESISTANCE TO ANTIMICROBIAL PEPTIDES
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: U32837; AAC23289.1; -.
 CC PIR: E64134; E64134.
 CC TIGR: H1642; -.
 CC InterPro: IPR003593; AAA ATPase.
 CC InterPro: IPR003439; ABC transporter.
 CC Pfam: PF00005; ABC_tran; 1.
 CC SMART: SM00382; AAA; 1.
 CC DR PROSITE: PS00211; ABC_TRANSPORTER_1; FALSE_NEG.
 CC DR PROSITE: PS50893; ABC_TRANSPORTER_2; 1.
 CC KW Peptide transport; Transport; Inner membrane; ATP-binding;
 CC Complete proteome.
 CC NP_BIND 45 52 ATP (POTENTIAL).
 CC SEQUENCE 269 AA; 30294 MW; E5C044EBCBDE801F CRC64;
 SQ
 Query Match 2.5%; Score 7; DB 1; Length 269;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 149 LARALIL 155
 DB 159 LARALIL 165

RP SEQUENCE FROM N.A.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
 RA Perry R.D.,
 RL "Genome sequence of Yersinia pestis KIM.",
 RL J. Bacteriol. 184:4601-4611(2002).
 CC -1- FUNCTION: This protein is a component of a D-methionine
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC
 CC -1- MICELLANEOUS: The meiniQ system is also to be able to transport
 CC the toxic methionine analog alpha-methyl-methionine (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the nipa lipoprotein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: A141416; CAC89914.1; -.
 CC EMBL: AE013911; AA86656.1; -.
 CC PIR: AG0131; AG0131.
 CC InterPro: IPR004872; Lipoprotein_9.
 CC InterPro: IPR000437; Prok_lipoprot_S.
 CC InterPro: IPR004478; YaeC.
 CC DR Pfam: PF03180; Lipoprotein_9; 1.
 CC DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 CC DR PROSITE: PS00013; PROKAR_LIPOPROTEIN_1.
 CC KW Transport; Amino-acid transport; Membrane; Lipoprotein; Signal;
 CC Complete proteome; Palmitate.
 CC FT SIGNAL 1 22 POTENTIAL.
 CC FT CHAIN 23 271 D-METHIONINE-BINDING LIPOPROTEIN METQ.
 CC FT LIPID 23 23 N-palmitoyl cysteine (Potential).
 CC FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
 CC SEQUENCE 271 AA; 29376 MW; 5214CEB0511F8C8 CRC64;
 SQ
 Query Match 2.5%; Score 7; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 160 GLIKLKD 166
 DB 153 GLIKLKD 159

RESULT 29
 ID METQ_YERPE STANDARD; PRT; 276 AA.
 AC Q08669; Q07364;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Outer membrane lipoprotein 2 precursor (PLP2).
 GN PLPB.
 OS Pasteurella haemolytica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 OX NCBI_TaxID=75985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=serotype A1;
 RX MEDLINE=9332810; PubMed=8335249;
 RA Murphy G.L., Whitworth L.C.;
 RA "Analysis of tandem, multiple genes encoding 30-kDa membrane proteins


```
RT in Pasteurella haemolytica A1."
RL Gene 129:107-111(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Serotype A1;
RX MEDLINE=94011378; PubMed=8406866;
RA Cooney B.J., Lo R.Y.C.;
RT "Three contiguous lipoprotein genes in Pasteurella haemolytica A1
RT which are homologous to a lipoprotein gene in Haemophilus influenzae
RT type b."
RL Infect. Immun. 61:4682-4688(1993).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -1- SIMILARITY: Belongs to the alpha lipoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L11037; AAA25539.1; -
DR EMBL, L16627; AAA25547.1; -
DR PIR, JN0752; JN0752.
DR InterPro: IPR004872; Lipoprotein 9.
DR InterPro: IPR000437; Prok_lipoprot_S.
DR InterPro: IPR004478; YaeC
DR Pfam: PF03180; Lipoprotein 9; 1.
DR TIGRfam: TIGR00363; TIGR00363; 1.
DR PROSITE: PS00013; PROCAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Multigene family; Palmitate.
FT SIGNAL 1 19
FT CHAIN 20 276
FT LIPID 20 20 N-palmitoyl cysteine (Probable).
FT LIPID 20 20 S-diacylglycerol cysteine (Probable).
FT CONFLICT 268 270 FNG -> LIVH (IN REF. 2).
SQ SEQUENCE 276 AA; 30157 MW; DF8F434ADD4D950 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 276;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGAT 139
DB 133 ELKDGAT 139

RESULT 30
NAK2_ECOL6 STANDARD; PRT; 291 AA.
AC Q8FDU8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative N-acetylmannosamine kinase 2 (EC 2.7.1.60) (MannAC kinase 2).
GN NANK2 OR C3638.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=O6:H / CFT073 / ATCC 700928;
RX MEDLINE=2388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G., Ili, Redford P., Roesch P.,
RA Rasko D., Buckles E.D., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Maslow G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
```

```
CC -1- FUNCTION: Catalyzes the phosphorylation of the N-acetylmannosamine
CC (MannAc) liberated from N-acetyl-neuraminic acid by the nanA
CC protein (Potential).
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-D-mannosamine = ADP + N-acetyl-
CC D-mannosamine 6-phosphate.
CC -1- PATHWAY: N-acetylmannosamine utilization.
CC -1- PATHWAY: N-acetylneuraminic acid (sialic acid) utilization.
CC -1- SIMILARITY: Belongs to the ROK (sugar kinases) family. Nank
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE016766; AAN82086.1; -
DR HAMAP, MF_01234; -; 1.
DR InterPro: IPR006000; ROK.
DR Pfam, PF00480; ROK; 1.
DR PROSITE, PS01125; ROK; 1.
KW Carbohydrate metabolism; Transferase; Kinase; ATP-binding;
KW Complete proteome.
FT NP BIND 5 12
FT NP BIND 132 139 ATP (POTENTIAL).
SQ SEQUENCE 291 AA; 30743 MW; 260BAD700BBB4E4 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 291;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 LAGSTK 127
DB 191 LAGSTK 197

RESULT 31
ADT1_BOVIN STANDARD; PRT; 297 AA.
AC P02722;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADP-ATP carrier protein, heart isoform T1 (ADP/ATP translocase 1)
DE (Adenine nucleotide translocator 1) (ANT 1).
GN SLC25A4 OR ANMT1
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228093; PubMed=2540808;
RA Powell S.J., Medd S.W., Runswick W.J., Walker J.E.;
RA "Two bovine genes for mitochondrial ADP/ATP translocase expressed
RT differences in various tissues."
RL Biochemistry 28:866-873(1989).
[2]
RP SEQUENCE.
RX MEDLINE=82188267; PubMed=7076130;
RA Aquila H., Mista D., Bullitz M., Klingenberg M.;
RT "Complete amino acid sequence of the ADP/ATP carrier from beef heart
RT mitochondria."
RL Hoppe-Seyler's Z. Physiol. Chem. 363:345-349(1982).
[3]
RP SEQUENCE OF 207-297 FROM N.A.
RX MEDLINE=86295775; PubMed=3017341;
RA Rasmussen U.B., Wohlbad H.;
RT "Bovine cardiac mitochondrial ADP/ATP-carrier: two distinct mRNAs and
```


RN J. Biol. Chem. 262:4355-4358 (1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chen C.N., Su Y., Baybayan P., Siruno A., Nagaraja R.,
 RA Mazarella R.A., Schlesinger D., Chen E.Y.,
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Becker M., Graves T., Ozersky P.,
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RX MEDLINE=23388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang U., Hsieh F.,
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebli T.B., Toshikiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.T., Abramson R.D., Miliak S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Raheij J., Helton E., Kettelman K., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [6]
 RP SEQUENCE OF 47-298 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=88124845; PubMed=2829183;
 RA Houdsworth J., Attardi G.,
 RT "Two distinct genes for ADP/ATP translocase are expressed at the mRNA
 RT level in adult human liver."
 RL Proc. Natl. Acad. Sci. U.S.A. 85:377-381 (1988).
 CC -1- FUNCTION: Catalyzes the exchange of ADP and ATP across the
 CC mitochondrial inner membrane.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
 CC -1- SIMILARITY: Contains 3 Solcar repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M57434; AAA51737.1; -
 DR EMBL: J02683; AAA55579.1; -
 DR EMBL: U78810; AAB39266.1; -
 DR EMBL: AC004000; AAB96347.1; -
 DR EMBL: BC056160; AAB56160.1; -
 DR EMBL: J03591; AAA36749.1; -
 DR PIR: A29132; A29132.
 DR GeneW: HGNC:10991; SLC25A5.
 DR MIM: 300150.
 DR GO: GO:000887; C: integral to plasma membrane; TAS.
 DR GO: GO:0015207; P: adenine transporter activity; TAS.
 DR GO: GO:006810; P: transport; TAS.
 DR InterPro: IPR002067; Mlt carrier.
 DR InterPro: IPR002030; Mlt uncoupling.
 DR InterPro: IPR001993; Mlt carrier.
 DR Gene: 254:57-66 (2000).

DR Pfam: PF00153; mltcarr. 3.
 DR PRINTS: PR00926; MITOCARRIER.
 DR PRINTS: PR00784; MITOCOUPLING.
 DR PROSITE: PS50920; SOLCAR; 3.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
 KW Multigene family.
 FT TRANSMEM 12 29 1 (POTENTIAL).
 FT TRANSMEM 73 91 2 (POTENTIAL).
 FT TRANSMEM 117 134 3 (POTENTIAL).
 FT TRANSMEM 176 195 4 (POTENTIAL).
 FT TRANSMEM 214 231 5 (POTENTIAL).
 FT TRANSMEM 273 291 6 (POTENTIAL).
 FT REPEAT 6 98 SOLCAR 1.
 FT REPEAT 111 201 SOLCAR 2.
 FT REPEAT 212 297 SOLCAR 3.
 FT CONFLICT 6 6 V -> L (IN REF. 2).
 FT CONFLICT 66 66 G -> E (IN REF. 2).
 FT CONFLICT 111 111 R -> L (IN REF. 4 AND 6).
 FT CONFLICT 162 162 V -> G (IN REF. 6).
 SQ SEQUENCE 298 AA; 32895 MW; F973C3AED92C49D3 CRC64;
 Query Match 2.5% Score 7; DB 1; Length 298;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 29 AAIKSTA 35
 DB 19 AAIKSTA 25
 RESULT 34
 ID ADP2_MOUSE STANDARD; PRT; 298 AA.
 AC P51881; O61311;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADP/ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2)
 DE (Adenine nucleotide translocator 2) (ANT 2).
 GN SLC25A5 OR ANT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=97059403; PubMed=8903724;
 RA Ellison J.W., Li X., Franke U., Shapiro L.J.,
 RT "Rapid evolution of human pseudautosomal genes and their mouse
 RT homologs."
 RL Mamm. Genome 7:25-30 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Sheldon J.G.;
 RL Thesis (1995), University of Cambridge, U.K.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RA Coester P., Laplace C.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP REVISIONS.
 RA Laplace C.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20432087; PubMed=10974536;
 RA Levy S.E., Chen Y.-S., Graham B.H., Wallace D.C.;
 RT "Expression and sequence analysis of the mouse adenine nucleotide
 RT translocase 1 and 2 genes."
 RL Gene 254:57-66 (2000).

ADT3_BOVIN STANDARD; PRT; 298 AA.
AC P32007;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE ADP,ATP carrier protein, isoform T2 (ADP/ATP translocase 3) (adenine nucleotide translocator 3) (ANT 3).
GN SLC25A6 OR ANT3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89229093; PubMed=2540808;
RA Powell S.J., Mead S.M., Runswick M.J., Walker J.E.;
RT "Two bovine genes for mitochondrial ADP/ATP translocase expressed differences in various tissues.";
RL Biochemistry 28:866-873(1989).
CC -1- FUNCTION: Catalyzes the exchange of ADP and ATP across the mitochondrial inner membrane.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.
CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
CC -1- SIMILARITY: Contains 3 Solcar repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24103; AAA30769.1; -
DR PIR; B43646; B43646.
DR InterPro; IPR002067; Mit_carrier.
DR InterPro; IPR002030; Mit_uncoupling.
DR InterPro; IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carri_3.
DR PRINTS; PR00926; MITOCARRIER.
DR PRINTS; PR00784; MTUNCOUPLING.
DR PROSITE; PS50920; SOLCAR_3.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport; MultiGene Family.
KM TRANSMEM 12 29 1 (POTENTIAL).
FT TRANSMEM 73 91 2 (POTENTIAL).
FT TRANSMEM 117 134 3 (POTENTIAL).
FT TRANSMEM 176 195 4 (POTENTIAL).
FT TRANSMEM 214 231 5 (POTENTIAL).
FT TRANSMEM 273 281 6 (POTENTIAL).
FT REPEAT 6 98 SOLCAR 1.
FT REPEAT 111 201 SOLCAR 2.
FT REPEAT 212 297 SOLCAR 3.
SQ SEQUENCE 298 AA; 32877 MW; 1C34E7DF6E84061 CRC64;
Query Match 2.5%; Score 7; DR 1; Length 298;
Best Local Similarity 100.0%; Fred No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ADT3_HUMAN STANDARD; PRT; 298 AA.
AC P12336; O96C49;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADP,ATP carrier protein, liver isoform T2 (ADP/ATP translocase 3) (adenine nucleotide translocator 3) (ANT 3).
GN SLC25A6 OR ANT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89236396; PubMed=2541251;
RA Cozens A.L., Runswick M.J., Walker J.E.;
RT "DNA sequences of two expressed nuclear genes for human mitochondrial ADP/ATP translocase.";
RL J. Mol. Biol. 206:261-280(1989).
CC [2]
CC SEQUENCE FROM N.A.
RA Zhou U., Yu W., Tang H., Mei G., Tsang Y.T.M., Bouck J., Gibbs R.A., Margolin J.F.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain, Cervix, Eye, and lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:11689-116903(2002).
RN [4]
RP SEQUENCE OF 36-298 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88124845; PubMed=2829183;
RA Houldsworth J., Altard G.;
RT "Two distinct genes for ADP/ATP translocase are expressed at the mRNA level in adult human liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:377-381(1988).
CC -1- FUNCTION: Catalyzes the exchange of ADP and ATP across the mitochondrial inner membrane.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.
CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
CC -1- SIMILARITY: Contains 3 Solcar repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03592; AAA36750.1; -
DR EMBL; AY007135; AAG01998.1; -
DR EMBL; BC007285; AA07295.1; -
DR EMBL; BC007850; AA07850.1; -

DR EMBL: BC008737; AAH08737.1; -
DR EMBL: BC008935; AAH08935.1; -
DR EMBL: BC014775; AAH14775.1; -
DR EMBL: BC01912; AAH1912.1; -
DR PIR: S03894; S03894.
DR Genem; HGNC:10992; SLC25A6.
DR MIM: 300151; -
DR MIM: 403000; -
DR GO: GO:000544; Mitochondrial inner membrane pre-sequence t. . .; TAS.
DR GO: GO:000547; P-ATP/ADP antiporter activity; TAS.
DR GO: GO:0006854; P-ATP/ADP exchange; TAS.
DR InterPro: IPR002067; Mt_carrier.
DR InterPro: IPR002030; Mitochondrial carrier.
DR Pfam: PF00153; mito_carr; 3-
DR PRINTS: PR00926; MITOCARRIER.
DR PRINTS: PR00784; MITOCARRIER.
DR PROSITE: PS00920; SOLCAR; 3.
DR KEGG: K01101; Mitochondrial inner membrane; Repeat; Transmembrane; Transport;
KW Multigene family.
FT TRANSMEM 12 29 1 (POTENTIAL).
FT TRANSMEM 73 91 2 (POTENTIAL).
FT TRANSMEM 117 134 3 (POTENTIAL).
FT TRANSMEM 176 195 4 (POTENTIAL).
FT TRANSMEM 214 231 5 (POTENTIAL).
FT TRANSMEM 273 291 6 (POTENTIAL).
FT REPEAT 6 98 SOLCAR 1.
FT REPEAT 111 201 SOLCAR 2.
FT REPEAT 112 297 SOLCAR 3.
FT REPEAT 105 108 SOLCAR 3.
FT CONFLICT 242 242 S -> F (IN REF. 3; AAH14775).
SQ SEQUENCE 298 AA; 32866 MW; 18534EP089672F CRC64;

Query Match 2.5%; Score 7; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAISKTA 35
DB 19 AAISKTA 25

RESULT 38
HEM3_NEIMA STANDARD; PRT; 311 AA.
ID HEM3_NEIMA
AC Q9UVS4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
synthase) (HMS) (Pre-uroporphyrinogen synthase).
GN HEMC OR NMA0718.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=2022556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,
Juels K., Leather S., Moule S., Mungall K., Quail M.A.,
Rajandream M.A., Rutherford K.W., Simmonds W., Skelton J.,
Srinivasan S., Spratt B.G., Barrall B.G.,
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491.";
RL Nature 404:502-506(2000).
CC -1- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
hydroxymethylbilane preuroporphyrinogen in several discrete steps.
CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
hydroxymethylbilane + 4 NH(3).
CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
hydroxymethylbilane + 4 NH(3).

CC -1- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
porphobilinogen subunits are added (By similarity).
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to the HMS family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL, restriction -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: AL162754; CAB84003.1; -
CC PIR: A81915; A81915.
CC HSSP: P06983; LPDA.
CC HAMAP: MF_00260; -; 1.
CC InterPro: IPR000860; Porphobil deam.
CC Pfam: PF01379; Porphobil deam.
CC Pfam: PF03900; Porphobil deam; 1.
CC PRINTS: PR00151; PORPHBDNASE.
CC ProDom: PD02745; Porphobil deam; 1.
CC TIGRFAMs: TIGR00212; hemc; 1.
CC PROSITE: PS00533; PORPHOBILINOGEN DEAM; 1.
KW Porphyrin biosynthesis; Transferase; Complete proteome.
FT BINDING 242 242 PYRROMETHANE COFACTOR (BY SIMILARITY).
SQ SEQUENCE 311 AA; 33448 MW; DFC7B5C6BFD3A03 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 NPKXLVI 186
DB 2 NPKXLVI 8

RESULT 39
HEM3_NEIMA STANDARD; PRT; 311 AA.
ID HEM3_NEIMA
AC Q9KOF6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
synthase) (HMS) (Pre-uroporphyrinogen synthase).
GN HEMC OR NMA0539.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Dougherty B.A.,
Hart D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
Cotton M.D., Uitterback T.R., Khouri H., Qin H., Vamathevan J.,
Gill J., Scarlato V., Nishigami V., Pizsa M., Grandi G., Sun L.,
Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.,
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
hydroxymethylbilane preuroporphyrinogen in several discrete steps.
CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
hydroxymethylbilane + 4 NH(3).
CC -1- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
porphobilinogen subunits are added (By similarity).

```
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to the HMBs family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002409; AAF40968.1; -.
CC PIR; B81188; B81188.
CC HSSP; P06983; 1PDA.
CC TIGR; NMB0533; -.
CC HAVAP; MF_00260; -; 1.
CC InterPro; IPR000860; Porphobil_deam.
CC Pfam; PF01379; Porphobil_deam; 1.
CC Pfam; PF03900; Porphobil_deam; 1.
CC PRINTS; PR00151; PORPHBDMASE.
CC PRODOM; PD002745; Porphobil_deam; 1.
CC TIGRFAMs; TIGR00212; hemc; 1.
CC PROSITE; PS00533; PORPHOBILINOGEN_DEAM; 1.
CC Porphyrin biosynthesis; Transferase; Complete proteome.
CC BINDING; 242 242 PYRROMETHANE COFACTOR (BY SIMILARITY).
CC SEQUENCE 311 AA; 33478 MW; 8446BDBD7BF0577F CRC64;
SQ
Query Match 2.5%; Score 7; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 NPKKLV 186
Db 2 NPKKLV 8
RESULT 40
TFBI_THEAC STANDARD; PRT; 312 AA.
ID TFBI_THEAC
AC Q9HJN7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription initiation factor TIB 1 (TFIIB 1).
GN TFBA OR TA0940.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruesp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RA "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RA acidophilum."
RT Nature 407:508-513(2000).
CC -1- FUNCTION: Stabilizes TBP binding to an archaeal box-A promoter.
CC Also responsible for recruiting RNA polymerase II to the pre-
CC initiation complex (DNA-TBP-TFIIB) (By similarity).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SIMILARITY: Belongs to the TFIIB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AL445066; CAC12069.1; -.
DR HSSP; P29095; 1A1S.
DR HAVAP; MF_00383; -; 1.
DR InterPro; IPR00670; Cyclin.
DR InterPro; IPR000812; TFIIB sub.
DR Pfam; PF00382; transcript_fac2; 2.
DR PRINTS; PR00685; TIFACTO1B.
DR SMART; SM00385; CYCLIN; 2.
DR PROSITE; PS00782; TFIIB; 2.
CC Transcription regulation; Repeat; Zinc-finger; Metal-binding; Zinc;
CC Complete proteome.
CC ZN_FING 16 38
CC REPEAT 129 212
CC METAL 223 304
CC METAL 16 16
CC METAL 19 16
CC METAL 35 35
CC METAL 38 38
CC SEQUENCE 312 AA; 35100 MW; 85BE5459CA526837 CRC64;
SQ
Query Match 2.5%; Score 7; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 AVAEVAG 56
Db 283 AVAEVAG 289
Search completed: June 16, 2004, 11:16:05
Job time : 29 secs
```

1000

1000

1000

1000

1000

1000

1000

Mon Jun 22 11:04:02 2004

us-10-018-672-2.oligo.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:17:18 ; Search time 48 Seconds
(without alignments)
1619.925 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 276
Sequence: 1 NMFGKINGICALASGIALAG.....TDEVEAKKQFKDGVIKGM 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1158786 seqs, 281726120 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1158786

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the chance being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	88.0	276	12	US-10-282-122A-63386 Sequence 63386, A
2	19	6.9	273	9	US-09-815-242-11067 Sequence 11067, A
3	19	6.9	273	12	US-10-282-122A-58240 Sequence 58240, A
4	19	6.9	273	12	US-10-380-817-2 Sequence 2, Appl1
5	19	6.9	273	12	US-10-380-817-4 Sequence 4, Appl1
6	19	6.9	273	12	US-10-380-817-6 Sequence 6, Appl1
7	19	6.9	273	12	US-10-380-817-8 Sequence 8, Appl1
8	19	6.9	273	12	US-10-380-817-10 Sequence 10, Appl1
9	16	5.8	271	12	US-10-282-122A-68391 Sequence 68391, A
10	13	4.7	270	12	US-10-282-122A-52838 Sequence 52838, A
11	12	4.3	276	12	US-10-282-122A-67387 Sequence 67387, A
12	11	4.0	67	13	US-10-092-243A-15387 Sequence 15, Appl1
13	11	4.0	271	12	US-10-282-122A-55926 Sequence 55926, A
14	11	4.0	271	12	US-10-282-122A-76195 Sequence 76195, A
15	10	3.6	276	12	US-10-282-122A-44792 Sequence 44792, A

16	9	3.3	241	12	US-10-282-122A-44788 Sequence 44788, A
17	9	3.3	256	12	US-10-282-122A-61127 Sequence 61127, A
18	9	3.3	262	12	US-10-282-122A-51371 Sequence 51371, A
19	9	3.3	270	12	US-10-282-122A-55804 Sequence 55804, A
20	9	3.3	273	12	US-10-282-122A-60558 Sequence 60558, A
21	9	3.3	275	12	US-10-282-122A-77174 Sequence 77174, A
22	8	2.9	96	9	US-09-764-853-805 Sequence 805, App
23	8	2.9	96	14	US-10-091-481-162 Sequence 162, App
24	8	2.9	143	12	US-10-424-559-262926 Sequence 262926, A
25	8	2.9	240	9	US-09-815-242-11656 Sequence 11656, A
26	8	2.9	263	12	US-10-282-122A-52976 Sequence 52976, A
27	8	2.9	265	12	US-10-282-122A-51201 Sequence 51201, A
28	8	2.9	270	12	US-10-282-122A-49339 Sequence 49339, A
29	8	2.9	271	9	US-09-815-242-10050 Sequence 10050, A
30	8	2.9	271	12	US-10-282-122A-43291 Sequence 43291, A
31	8	2.9	271	12	US-10-282-122A-59600 Sequence 59600, A
32	8	2.9	281	12	US-10-282-122A-74389 Sequence 74389, A
33	8	2.9	282	12	US-10-282-122A-57911 Sequence 57911, A
34	8	2.9	295	12	US-10-282-122A-47704 Sequence 47704, A
35	8	2.9	379	9	US-09-764-853-621 Sequence 621, App
36	8	2.9	467	15	US-10-369-493-20865 Sequence 20865, A
37	8	2.9	641	15	US-10-369-493-15592 Sequence 15592, A
38	8	2.9	739	9	US-09-854-845-45 Sequence 45, Appl1
39	8	2.9	744	9	US-09-854-845-43 Sequence 43, Appl1
40	8	2.9	766	9	US-09-854-845-49 Sequence 49, Appl1
41	8	2.9	771	9	US-09-854-845-47 Sequence 47, Appl1
42	8	2.9	838	9	US-09-854-845-23 Sequence 23, Appl1
43	8	2.9	838	15	US-10-150-115-49 Sequence 49, Appl1
44	8	2.9	838	15	US-10-369-072-19 Sequence 19, Appl1
45	8	2.9	843	9	US-09-854-845-27 Sequence 27, Appl1
46	8	2.9	865	9	US-09-854-845-33 Sequence 33, Appl1
47	8	2.9	870	9	US-09-854-845-31 Sequence 31, Appl1
48	7	2.5	69	12	US-10-424-559-195672 Sequence 195672, A
49	7	2.5	69	12	US-10-424-559-200956 Sequence 200956, A
50	7	2.5	73	12	US-10-335-977-4881 Sequence 4881, App
51	7	2.5	85	12	US-10-424-559-271432 Sequence 271432, A
52	7	2.5	87	12	US-10-424-559-271850 Sequence 271850, A
53	7	2.5	92	12	US-10-296-115-1035 Sequence 1035, App
54	7	2.5	94	12	US-10-424-559-212355 Sequence 212355, A
55	7	2.5	131	9	US-09-925-300-1808 Sequence 1808, App
56	7	2.5	132	9	US-09-925-301-1459 Sequence 1459, App
57	7	2.5	133	14	US-10-259-642-12 Sequence 12, Appl1
58	7	2.5	134	12	US-10-424-559-163134 Sequence 163134, A
59	7	2.5	146	12	US-10-424-559-167109 Sequence 167109, A
60	6	2.5	153	12	US-10-335-977-9642 Sequence 9642, App
61	6	2.5	154	14	US-10-299-642-14 Sequence 14, Appl1
62	6	2.5	156	12	US-10-424-559-205606 Sequence 205606, A
63	6	2.5	158	12	US-10-425-114-53426 Sequence 53426, A
64	6	2.5	159	9	US-09-963-791-18 Sequence 18, Appl1
65	6	2.5	159	12	US-10-419-276-18 Sequence 18, Appl1
66	6	2.5	174	14	US-10-017-161-700 Sequence 700, App
67	6	2.5	174	15	US-10-292-798-612 Sequence 612, App
68	6	2.5	178	15	US-10-292-798-606 Sequence 606, App
69	7	2.5	191	12	US-10-424-559-210718 Sequence 210718, A
70	7	2.5	194	9	US-09-864-761-33808 Sequence 33808, A
71	7	2.5	204	12	US-10-424-559-203852 Sequence 203852, A
72	7	2.5	220	12	US-10-282-122A-69434 Sequence 69434, A
73	7	2.5	228	15	US-10-264-049-4337 Sequence 4337, App
74	7	2.5	234	12	US-10-335-977-8107 Sequence 8107, App
75	7	2.5	235	12	US-10-335-977-8108 Sequence 8108, App
76	7	2.5	238	12	US-10-425-114-69447 Sequence 69447, A
77	7	2.5	244	12	US-10-450-055-38 Sequence 38, Appl1
78	7	2.5	247	12	US-10-282-122A-54377 Sequence 54377, A
79	7	2.5	256	12	US-10-282-122A-67627 Sequence 67627, A
80	7	2.5	260	9	US-09-815-242-12110 Sequence 12110, A
81	7	2.5	261	12	US-10-282-122A-66844 Sequence 66844, A
82	7	2.5	262	12	US-10-282-122A-71312 Sequence 71312, A
83	7	2.5	262	12	US-10-282-122A-46629 Sequence 46629, A
84	7	2.5	262	12	US-10-282-122A-77131 Sequence 77131, A
85	7	2.5	265	12	US-10-282-122A-67132 Sequence 67132, A
86	7	2.5	268	12	US-10-282-122A-42798 Sequence 42798, A
87	7	2.5	268	12	US-10-282-122A-49256 Sequence 49256, A
88	7	2.5	268	12	US-10-282-122A-55568 Sequence 55568, A

89	2.5	268	12	US-10-282-122A-73574	Sequence 73574, A	162	7	2.5	475	14	US-10-299-642-24	Sequence 24, Appl
90	2.5	268	12	US-10-282-122A-74991	Sequence 74991, A	163	7	2.5	477	11	US-09-826-509-493	Sequence 493, Appl
91	2.5	268	12	US-10-282-122A-76001	Sequence 76001, A	164	7	2.5	477	12	US-09-964-956-47	Sequence 47, Appl
92	2.5	269	12	US-10-282-122A-58568	Sequence 58568, A	165	7	2.5	477	14	US-10-377-078-2	Sequence 2, Appl
93	2.5	269	12	US-10-282-122A-59439	Sequence 59439, A	166	7	2.5	477	14	US-10-377-078-3	Sequence 3, Appl
94	2.5	269	12	US-10-282-122A-68970	Sequence 68970, A	167	7	2.5	477	14	US-10-425-567A-100	Sequence 100, Appl
95	2.5	270	12	US-10-282-122A-45531	Sequence 45531, A	168	7	2.5	477	14	US-10-299-642-8	Sequence 8, Appl
96	2.5	271	9	US-09-815-242-11460	Sequence 11460, A	169	7	2.5	477	15	US-10-352-684A-14	Sequence 14, Appl
97	2.5	271	9	US-09-815-242-11624	Sequence 11624, A	170	7	2.5	477	15	US-09-964-956-15	Sequence 15, Appl
98	2.5	271	12	US-10-282-122A-59034	Sequence 59034, A	171	7	2.5	486	12	US-09-964-956-50	Sequence 50, Appl
99	2.5	271	12	US-10-282-122A-71343	Sequence 71343, A	172	7	2.5	486	12	US-10-282-122A-70987	Sequence 8, Appl
100	2.5	271	12	US-10-282-122A-78174	Sequence 78174, A	173	7	2.5	496	12	US-10-352-684A-14	Sequence 14, Appl
101	2.5	271	12	US-10-282-122A-78367	Sequence 78367, A	174	7	2.5	499	12	US-10-282-122A-41748	Sequence 12, Appl
102	2.5	271	12	US-10-335-977-4882	Sequence 4882, Ap	175	7	2.5	507	9	US-09-963-791-10	Sequence 10, Appl
103	2.5	271	12	US-10-335-977-4883	Sequence 4883, Ap	176	7	2.5	507	12	US-10-419-276-10	Sequence 10, Appl
104	2.5	272	12	US-10-282-122A-52590	Sequence 52590, A	177	7	2.5	556	14	US-10-335-985-7533	Sequence 7533, Ap
105	2.5	274	12	US-10-282-122A-77956	Sequence 77956, A	178	7	2.5	580	15	US-10-369-493-2839	Sequence 2839, Ap
106	2.5	275	14	US-10-156-761-7735	Sequence 7735, Ap	179	7	2.5	589	9	US-09-963-791-12	Sequence 12, Appl
107	2.5	276	12	US-10-282-122A-60923	Sequence 60923, A	180	7	2.5	589	12	US-10-419-276-12	Sequence 12, Appl
108	2.5	280	12	US-10-335-977-4884	Sequence 4884, Ap	181	7	2.5	591	12	US-10-425-114-63542	Sequence 63542, A
109	2.5	280	15	US-10-369-493-9042	Sequence 9042, Ap	182	7	2.5	633	12	US-10-282-122A-34071	Sequence 13177, A
110	2.5	285	12	US-10-425-114-57973	Sequence 57973, A	183	7	2.5	652	9	US-09-815-242-13673	Sequence 13673, A
111	2.5	289	14	US-10-156-761-8987	Sequence 8987, Ap	184	7	2.5	652	12	US-10-282-122A-73973	Sequence 73973, A
112	2.5	298	9	US-09-811-094-32	Sequence 32, Appl	185	7	2.5	670	12	US-10-282-122A-47105	Sequence 47105, A
113	2.5	298	9	US-09-811-094-33	Sequence 33, Appl	186	7	2.5	691	8	US-08-834-666A-12	Sequence 12, Appl
114	2.5	298	9	US-09-810-644-32	Sequence 32, Appl	187	7	2.5	691	8	US-08-834-666A-22	Sequence 22, Appl
115	2.5	298	9	US-09-810-644-33	Sequence 33, Appl	188	7	2.5	692	9	US-09-738-626-5979	Sequence 6979, Ap
116	2.5	298	9	US-09-185-904A-32	Sequence 32, Appl	189	7	2.5	757	9	US-09-963-791-24	Sequence 24, Appl
117	2.5	298	9	US-09-185-904A-32	Sequence 32, Appl	190	7	2.5	757	12	US-10-419-276-24	Sequence 24, Appl
118	2.5	298	12	US-10-403-571-154	Sequence 154, App	191	7	2.5	852	12	US-10-282-122A-65071	Sequence 65071, A
119	2.5	298	16	US-10-408-765A-1531	Sequence 1531, Ap	192	7	2.5	852	12	US-10-282-122A-65901	Sequence 65901, A
120	2.5	298	16	US-10-408-765A-2398	Sequence 2398, Ap	193	7	2.5	859	9	US-09-918-112A-5	Sequence 5, Appl
121	2.5	299	12	US-10-424-599-280669	Sequence 280669, A	194	7	2.5	860	9	US-09-738-626-464	Sequence 464, Ap
122	2.5	300	15	US-10-369-493-5088	Sequence 5088, Ap	195	7	2.5	860	9	US-09-981-151A-34	Sequence 34, Appl
123	2.5	309	14	US-10-304-928-10	Sequence 10, Appl	196	7	2.5	875	12	US-10-282-122A-60678	Sequence 60678, A
124	2.5	310	9	US-09-963-791-8	Sequence 8, Appl	197	7	2.5	908	9	US-09-963-791-2	Sequence 2, Appl
125	2.5	310	12	US-10-419-276-8	Sequence 8, Appl	198	7	2.5	908	12	US-10-419-276-2	Sequence 2, Appl
126	2.5	311	12	US-10-282-122A-65475	Sequence 65475, A	199	7	2.5	959	9	US-09-788-043C-1	Sequence 1, Appl
127	2.5	311	12	US-10-282-122A-65779	Sequence 65779, A	200	7	2.5	1036	12	US-10-282-122A-76570	Sequence 76570, A
128	2.5	312	15	US-10-369-493-13581	Sequence 13581, A	201	7	2.5	1120	16	US-10-467-042-6	Sequence 6, Appl
129	2.5	312	15	US-10-369-493-18216	Sequence 18216, A	202	7	2.5	1120	16	US-10-282-122A-69138	Sequence 580, Ap
130	2.5	317	9	US-09-963-791-16	Sequence 16, Appl	203	7	2.5	1130	15	US-10-369-493-5580	Sequence 5580, Ap
131	2.5	317	12	US-10-419-276-16	Sequence 16, Appl	204	7	2.5	1130	15	US-10-332-565-7014	Sequence 7014, Ap
132	2.5	332	15	US-10-469-493-8019	Sequence 8019, Ap	205	7	2.5	1178	12	US-10-282-122A-5234	Sequence 5234, A
133	2.5	337	10	US-09-769-734-29	Sequence 29, Appl	206	7	2.5	1233	9	US-09-738-626-4312	Sequence 4312, Ap
134	2.5	337	12	US-10-107-431-53	Sequence 53, Appl	207	7	2.5	1488	10	US-09-940-168-21	Sequence 21, Appl
135	2.5	341	15	US-10-369-493-641	Sequence 641, App	208	7	2.5	1509	10	US-09-940-168-23	Sequence 23, Appl
136	2.5	347	9	US-09-977-577-3	Sequence 3, Appl	209	7	2.5	1517	10	US-09-940-168-19	Sequence 19, Appl
137	2.5	352	9	US-09-815-242-10053	Sequence 10053, A	210	7	2.5	1517	10	US-09-940-168-19	Sequence 71, Appl
138	2.5	355	14	US-10-017-161-694	Sequence 694, App	211	7	2.5	1517	10	US-10-376-537-12	Sequence 72, Appl
139	2.5	355	14	US-10-017-161-722	Sequence 722, App	212	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
140	2.5	355	15	US-10-292-798-632	Sequence 632, App	213	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
141	2.5	356	9	US-09-963-791-20	Sequence 20, Appl	214	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
142	2.5	356	12	US-10-419-276-20	Sequence 20, Appl	215	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
143	2.5	376	12	US-10-425-114-55210	Sequence 55210, A	216	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
144	2.5	380	12	US-10-092-771-6	Sequence 6, Appl	217	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
145	2.5	386	12	US-10-424-599-21898	Sequence 21898, A	218	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
146	2.5	411	14	US-10-350-298-1	Sequence 1, Appl	219	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
147	2.5	419	9	US-09-712-363-489	Sequence 289, App	220	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
148	2.5	419	12	US-10-282-122A-62797	Sequence 62797, A	221	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
149	2.5	419	12	US-10-282-122A-64940	Sequence 64940, A	222	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
150	2.5	424	9	US-09-738-626-4306	Sequence 4306, Ap	223	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
151	2.5	438	12	US-09-963-791-32	Sequence 22, Appl	224	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
152	2.5	438	12	US-10-419-276-22	Sequence 22, Appl	225	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
153	2.5	457	12	US-09-964-956-42	Sequence 49, Appl	226	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
154	2.5	458	12	US-09-964-956-49	Sequence 51, Appl	227	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
155	2.5	460	9	US-09-964-956-51	Sequence 51, Appl	228	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
156	2.5	468	12	US-09-738-626-4223	Sequence 4223, Ap	229	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
157	2.5	468	12	US-09-963-791-6	Sequence 6, Appl	230	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
158	2.5	468	12	US-10-419-276-6	Sequence 6, Appl	231	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
159	2.5	468	15	US-10-369-493-6465	Sequence 6465, Ap	232	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
160	2.5	471	15	US-10-369-493-21077	Sequence 21077, A	233	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl
161	2.5	475	12	US-09-964-956-48	Sequence 1456, Ap	234	7	2.5	1517	10	US-10-376-537-21	Sequence 21, Appl

235	6	2.2	29	16	US-10-253-493-1431	Sequence 1431, Ap	308	6	2.2	103	12	US-10-424-599-237711	Sequence 237711, Ap
236	6	2.2	31	13	US-10-032-330-38	Sequence 38, Appl	309	6	2.2	103	14	US-10-032-330-38	Sequence 7785, Ap
237	6	2.2	35	13	US-10-016-634A-132	Sequence 132, Appl	310	6	2.2	105	12	US-10-424-599-238378	Sequence 238378, Ap
238	6	2.2	40	12	US-10-424-599-153611	Sequence 153611, Ap	311	6	2.2	105	12	US-10-424-599-153611	Sequence 11, Appl
239	6	2.2	43	12	US-10-424-599-1865934	Sequence 1865934, Ap	312	6	2.2	105	15	US-10-074-978A-140	Sequence 140, Appl
240	6	2.2	43	14	US-10-321-857-136	Sequence 136, Appl	313	6	2.2	107	12	US-10-424-599-253985	Sequence 253985, Ap
241	6	2.2	43	14	US-10-318-675-136	Sequence 136, Appl	314	6	2.2	109	12	US-10-424-599-212413	Sequence 212413, Ap
242	6	2.2	45	14	US-10-029-386-28062	Sequence 28062, A	315	6	2.2	111	9	US-09-738-628-3654	Sequence 3654, Ap
243	6	2.2	50	12	US-10-424-599-239203	Sequence 239203, Ap	316	6	2.2	111	11	US-09-864-408A-170	Sequence 170, Appl
244	6	2.2	51	12	US-10-424-599-182141	Sequence 182141, Ap	317	6	2.2	113	9	US-10-424-599-195737	Sequence 195737, Ap
245	6	2.2	52	10	US-09-764-872-261	Sequence 261, Appl	318	6	2.2	114	12	US-09-811-268-149	Sequence 149, Appl
246	6	2.2	52	12	US-10-424-599-226535	Sequence 226535, Ap	319	6	2.2	116	12	US-10-282-122A-73031	Sequence 73031, A
247	6	2.2	52	12	US-10-424-599-240019	Sequence 240019, Ap	320	6	2.2	116	12	US-10-424-599-198173	Sequence 198173, Ap
248	6	2.2	53	12	US-10-424-599-225432	Sequence 225432, Ap	321	6	2.2	117	11	US-09-981-151A-79	Sequence 79, Appl
249	6	2.2	54	9	US-09-939-980-320	Sequence 320, Appl	322	6	2.2	117	11	US-09-864-408A-836	Sequence 836, Ap
250	6	2.2	55	12	US-10-424-599-279936	Sequence 279936, Ap	323	6	2.2	117	12	US-10-424-599-170868	Sequence 170868, Ap
251	6	2.2	56	12	US-10-424-599-167091	Sequence 167091, Ap	324	6	2.2	119	9	US-09-858-081-4	Sequence 4, Appl
252	6	2.2	57	12	US-10-424-599-152661	Sequence 152661, Ap	325	6	2.2	119	9	US-09-858-068-4	Sequence 4, Appl
253	6	2.2	58	12	US-10-424-599-156664	Sequence 156664, Ap	326	6	2.2	119	12	US-09-961-656-7	Sequence 7, Appl
254	6	2.2	58	14	US-10-029-386-31026	Sequence 31026, A	327	6	2.2	119	13	US-10-014-070-7	Sequence 7, Appl
255	6	2.2	60	12	US-10-424-599-244607	Sequence 244607, Ap	328	6	2.2	119	13	US-10-153-316-4	Sequence 4, Appl
256	6	2.2	62	12	US-10-424-599-148564	Sequence 148564, Ap	329	6	2.2	119	14	US-10-103-377C-4	Sequence 4, Appl
257	6	2.2	63	15	US-10-435-766-78	Sequence 78, Appl	330	6	2.2	119	15	US-10-028-248A-85	Sequence 85, Appl
258	6	2.2	64	12	US-10-280-139-10	Sequence 10, Appl	331	6	2.2	119	15	US-10-107-782-85	Sequence 85, Appl
259	6	2.2	65	12	US-10-424-599-283916	Sequence 283916, Ap	332	6	2.2	120	12	US-10-424-599-255944	Sequence 255944, Ap
260	6	2.2	66	10	US-09-983-802-654	Sequence 654, Appl	333	6	2.2	120	12	US-10-425-114-62096	Sequence 11462096, Ap
261	6	2.2	66	12	US-09-973-278-372	Sequence 278372, Ap	334	6	2.2	121	12	US-10-424-599-174541	Sequence 174541, Ap
262	6	2.2	66	12	US-10-424-599-234063	Sequence 234063, Ap	335	6	2.2	121	12	US-10-424-599-175317	Sequence 175317, Ap
263	6	2.2	66	12	US-10-424-599-256782	Sequence 256782, Ap	336	6	2.2	121	12	US-10-425-114-61185	Sequence 11461185, Ap
264	6	2.2	66	12	US-10-424-599-277312	Sequence 277312, Ap	337	6	2.2	121	15	US-10-264-237-2563	Sequence 2563, Ap
265	6	2.2	66	12	US-10-424-599-285501	Sequence 285501, Ap	338	6	2.2	123	9	US-09-864-761-18806	Sequence 48806, A
266	6	2.2	66	12	US-09-984-490-654	Sequence 654, Appl	339	6	2.2	123	9	US-10-424-599-142864	Sequence 142864, Ap
267	6	2.2	67	12	US-10-424-599-214538	Sequence 214538, Ap	340	6	2.2	123	12	US-10-424-599-178133	Sequence 178133, Ap
268	6	2.2	67	12	US-10-424-599-262099	Sequence 262099, Ap	341	6	2.2	123	12	US-10-424-599-247128	Sequence 247128, Ap
269	6	2.2	68	12	US-10-424-599-213493	Sequence 213493, Ap	342	6	2.2	124	12	US-10-424-599-244445	Sequence 244445, Ap
270	6	2.2	68	12	US-10-424-599-281467	Sequence 281467, Ap	343	6	2.2	126	12	US-10-424-599-247128	Sequence 247128, Ap
271	6	2.2	72	12	US-10-424-599-183635	Sequence 183635, Ap	344	6	2.2	127	12	US-10-424-599-215154	Sequence 215154, Ap
272	6	2.2	72	12	US-10-424-599-236659	Sequence 236659, Ap	345	6	2.2	127	14	US-10-156-761-11456	Sequence 11456, A
273	6	2.2	73	12	US-10-424-599-168202	Sequence 168202, Ap	346	6	2.2	128	12	US-10-282-122A-51178	Sequence 51178, A
274	6	2.2	74	14	US-10-424-599-283720	Sequence 283720, Ap	347	6	2.2	128	12	US-10-282-122A-51978	Sequence 51978, A
275	6	2.2	74	14	US-10-029-386-31524	Sequence 31524, A	348	6	2.2	129	12	US-10-424-599-226333	Sequence 226333, Ap
276	6	2.2	75	12	US-10-424-599-154202	Sequence 154202, Ap	349	6	2.2	131	10	US-09-903-190-175	Sequence 190175, Ap
277	6	2.2	76	12	US-10-424-599-179447	Sequence 179447, Ap	350	6	2.2	131	16	US-10-461-990-30	Sequence 30, Appl
278	6	2.2	80	12	US-10-424-599-229412	Sequence 229412, Ap	351	6	2.2	132	11	US-09-864-408A-2114	Sequence 2114, Ap
279	6	2.2	81	12	US-10-282-122A-52594	Sequence 52594, A	352	6	2.2	132	15	US-10-424-599-240909	Sequence 240909, Ap
280	6	2.2	81	12	US-10-424-599-160522	Sequence 160522, Ap	353	6	2.2	132	15	US-10-094-886-20	Sequence 20, Appl
281	6	2.2	82	10	US-09-903-463-16	Sequence 16, Appl	354	6	2.2	133	14	US-10-195-730-140	Sequence 730140, Ap
282	6	2.2	83	12	US-10-424-599-267319	Sequence 267319, Ap	355	6	2.2	133	14	US-09-833-245-170	Sequence 245170, Ap
283	6	2.2	84	12	US-10-282-122A-60406	Sequence 60406, A	356	6	2.2	135	12	US-10-424-599-184282	Sequence 184282, Ap
284	6	2.2	84	12	US-10-424-599-194766	Sequence 194766, Ap	357	6	2.2	135	12	US-10-424-599-276407	Sequence 276407, Ap
285	6	2.2	84	15	US-10-074-978A-291	Sequence 291, Appl	358	6	2.2	136	12	US-10-424-599-148636	Sequence 148636, Ap
286	6	2.2	85	12	US-10-424-599-164049	Sequence 164049, Ap	359	6	2.2	136	12	US-10-424-599-153334	Sequence 153334, Ap
287	6	2.2	85	12	US-10-424-599-267982	Sequence 267982, Ap	360	6	2.2	138	12	US-10-424-599-239007	Sequence 239007, Ap
288	6	2.2	85	12	US-10-424-599-280762	Sequence 280762, Ap	361	6	2.2	139	12	US-10-424-599-276923	Sequence 276923, Ap
289	6	2.2	86	12	US-10-424-599-143476	Sequence 143476, Ap	362	6	2.2	139	12	US-10-627-676-574	Sequence 676574, Ap
290	6	2.2	87	9	US-09-764-864-1130	Sequence 1130, Ap	363	6	2.2	139	12	US-10-450-055-44	Sequence 574, Appl
291	6	2.2	87	12	US-10-424-599-228388	Sequence 228388, Ap	364	6	2.2	140	15	US-10-108-860A-2564	Sequence 2564, Ap
292	6	2.2	88	15	US-10-264-049-3645	Sequence 3645, Ap	365	6	2.2	141	15	US-10-425-114-68389	Sequence 68389, Ap
293	6	2.2	89	9	US-09-764-864-1545	Sequence 1545, Ap	366	6	2.2	141	15	US-10-074-978A-14	Sequence 14, Appl
294	6	2.2	89	12	US-10-424-599-171497	Sequence 171497, Ap	367	6	2.2	142	12	US-10-424-599-205163	Sequence 205163, Ap
295	6	2.2	89	12	US-10-424-599-173898	Sequence 173898, Ap	368	6	2.2	143	12	US-10-424-599-198889	Sequence 198889, Ap
296	6	2.2	90	12	US-10-424-599-185961	Sequence 185961, Ap	369	6	2.2	144	12	US-10-282-122A-76111	Sequence 76111, A
297	6	2.2	90	12	US-10-424-599-283528	Sequence 283528, Ap	370	6	2.2	144	12	US-10-282-122A-76111	Sequence 76111, A
298	6	2.2	91	12	US-10-424-599-283528	Sequence 283528, Ap	371	6	2.2	144	12	US-10-424-599-149499	Sequence 149499, Ap
299	6	2.2	93	12	US-10-424-599-180542	Sequence 180542, Ap	372	6	2.2	145	12	US-10-282-122A-59117	Sequence 59117, A
300	6	2.2	93	12	US-10-424-599-211048	Sequence 211048, Ap	373	6	2.2	146	14	US-10-091-007-196	Sequence 196, Appl
301	6	2.2	93	15	US-10-074-978A-142	Sequence 142, Appl	374	6	2.2	147	9	US-09-895-913A-346	Sequence 346, Appl
302	6	2.2	94	12	US-10-424-599-279209	Sequence 279209, Ap	375	6	2.2	147	12	US-10-282-122A-54653	Sequence 54653, A
303	6	2.2	96	12	US-10-424-599-228371	Sequence 228371, Ap	376	6	2.2	148	12	US-10-282-122A-60619	Sequence 60619, A
304	6	2.2	97	12	US-10-424-599-265936	Sequence 265936, Ap	377	6	2.2	148	12	US-10-282-122A-67938	Sequence 67938, A
305	6	2.2	99	11	US-09-864-408A-1178	Sequence 1178, Ap	378	6	2.2	148	12	US-10-282-122A-65305	Sequence 65305, A
306	6	2.2	100	15	US-10-282-122A-63523	Sequence 63523, A	379	6	2.2	149	12	US-10-282-122A-42921	Sequence 42921, A
307	6	2.2	102	12	US-10-108-260A-4375	Sequence 4375, Ap	380	6	2.2	149	12	US-10-424-599-207991	Sequence 207991, Ap

381	6	2.2	150	10	US-09-764-891-4480	Sequence 4480, Ap	454	6	2.2	166	13	US-10-051-643-109	Sequence 109, App
382	6	2.2	151	12	US-10-282-122A-62416	Sequence 62416, A	455	6	2.2	166	14	US-10-205-979-36	Sequence 36, App1
383	6	2.2	151	12	US-10-282-122A-64879	Sequence 64879, A	456	6	2.2	166	15	US-10-389-674-38	Sequence 38, App1
384	6	2.2	151	13	US-10-062-254-52	Sequence 52, App1	457	6	2.2	166	15	US-10-389-674-42	Sequence 42, App1
385	6	2.2	151	15	US-10-094-886-18	Sequence 18, App1	458	6	2.2	166	15	US-10-389-674-43	Sequence 43, App1
386	6	2.2	153	12	US-10-282-122A-63692	Sequence 63692, A	459	6	2.2	166	15	US-10-389-674-44	Sequence 44, App1
387	6	2.2	153	14	US-10-091-007-146	Sequence 146, App	460	6	2.2	166	15	US-10-389-674-46	Sequence 46, App1
388	6	2.2	154	12	US-09-726-643-77	Sequence 77, App1	461	6	2.2	166	15	US-10-389-674-48	Sequence 48, App1
389	6	2.2	154	12	US-10-016-248-36	Sequence 36, App1	462	6	2.2	166	15	US-10-389-674-50	Sequence 50, App1
390	6	2.2	154	13	US-10-042-141-77	Sequence 77, App1	463	6	2.2	166	15	US-10-389-674-51	Sequence 51, App1
391	6	2.2	154	14	US-10-220-289-2	Sequence 2, App1	464	6	2.2	166	15	US-10-389-674-55	Sequence 55, App1
392	6	2.2	154	14	US-10-156-761-9045	Sequence 9045, Ap	465	6	2.2	166	15	US-10-389-674-56	Sequence 56, App1
393	6	2.2	155	9	US-09-925-301-1561	Sequence 1561, Ap	466	6	2.2	166	15	US-10-389-674-57	Sequence 57, App1
394	6	2.2	155	9	US-09-738-973-114	Sequence 114, App	467	6	2.2	166	15	US-10-389-674-59	Sequence 59, App1
395	6	2.2	155	9	US-09-854-133-114	Sequence 114, App	468	6	2.2	166	15	US-10-389-674-60	Sequence 60, App1
396	6	2.2	155	12	US-10-335-977-8128	Sequence 8128, Ap	469	6	2.2	166	15	US-10-389-674-61	Sequence 61, App1
397	6	2.2	155	12	US-09-815-242-12062	Sequence 114, App	470	6	2.2	166	15	US-10-389-674-62	Sequence 62, App1
398	6	2.2	156	9	US-09-144-649A-114	Sequence 114, App	471	6	2.2	166	15	US-10-389-674-63	Sequence 63, App1
399	6	2.2	156	12	US-09-764-891-3400	Sequence 3400, Ap	472	6	2.2	166	15	US-10-389-674-64	Sequence 64, App1
400	6	2.2	156	12	US-10-282-122A-66739	Sequence 66739, A	473	6	2.2	166	15	US-10-389-674-65	Sequence 65, App1
401	6	2.2	156	12	US-10-424-599-188143	Sequence 188143, A	474	6	2.2	166	15	US-10-389-674-67	Sequence 67, App1
402	6	2.2	156	12	US-10-424-599-188143	Sequence 188143, A	475	6	2.2	166	15	US-10-389-674-69	Sequence 69, App1
403	6	2.2	157	14	US-10-156-761-9744	Sequence 9744, Ap	476	6	2.2	166	15	US-10-389-674-70	Sequence 70, App1
404	6	2.2	157	16	US-10-415-969-40	Sequence 40, App1	477	6	2.2	166	15	US-10-389-674-79	Sequence 79, App1
405	6	2.2	158	9	US-09-738-626-6824	Sequence 6824, Ap	478	6	2.2	166	15	US-10-389-674-80	Sequence 80, App1
406	6	2.2	158	12	US-10-424-599-178819	Sequence 178819, Ap	479	6	2.2	166	15	US-10-389-674-84	Sequence 84, App1
407	6	2.2	158	12	US-10-335-977-8129	Sequence 8129, Ap	480	6	2.2	166	15	US-10-389-674-84	Sequence 84, App1
408	6	2.2	159	10	US-09-862-540-47	Sequence 8129, Ap	481	6	2.2	166	15	US-10-615-723-9	Sequence 9, App1
409	6	2.2	159	12	US-10-423-114-5572	Sequence 5572, A	482	6	2.2	166	15	US-10-615-723-11	Sequence 11, App1
410	6	2.2	159	15	US-10-369-493-16130	Sequence 16130, A	483	6	2.2	166	15	US-10-615-723-30	Sequence 30, App1
411	6	2.2	160	12	US-10-424-599-233206	Sequence 233206, A	484	6	2.2	166	15	US-10-615-723-32	Sequence 32, App1
412	6	2.2	160	12	US-10-425-114-62350	Sequence 62350, A	485	6	2.2	166	15	US-10-615-723-36	Sequence 36, App1
413	6	2.2	160	12	US-10-425-114-66329	Sequence 66329, A	486	6	2.2	166	15	US-10-615-723-38	Sequence 38, App1
414	6	2.2	161	9	US-09-726-643-78	Sequence 78, App1	487	6	2.2	166	15	US-10-615-723-38	Sequence 38, App1
415	6	2.2	161	10	US-09-997-672-9	Sequence 9, App1	488	6	2.2	166	15	US-10-615-723-42	Sequence 42, App1
416	6	2.2	161	12	US-10-424-599-234753	Sequence 234753, A	489	6	2.2	166	15	US-10-615-723-42	Sequence 42, App1
417	6	2.2	161	13	US-10-042-141-78	Sequence 78, App1	490	6	2.2	166	15	US-10-415-869-30	Sequence 30, App1
418	6	2.2	162	12	US-10-424-599-275145	Sequence 275145, A	491	6	2.2	166	15	US-10-415-869-74	Sequence 74, App1
419	6	2.2	162	16	US-10-461-990-31	Sequence 31, App1	492	6	2.2	166	15	US-10-415-869-78	Sequence 78, App1
420	6	2.2	163	12	US-10-424-599-156728	Sequence 156728, A	493	6	2.2	166	15	US-10-415-869-80	Sequence 80, App1
421	6	2.2	164	9	US-09-950-313-1	Sequence 3, App1	494	6	2.2	166	15	US-10-415-869-82	Sequence 82, App1
422	6	2.2	164	9	US-09-950-313-3	Sequence 3, App1	495	6	2.2	166	15	US-10-415-869-84	Sequence 84, App1
423	6	2.2	165	12	US-10-282-122A-44521	Sequence 44521, A	496	6	2.2	166	15	US-10-415-869-86	Sequence 86, App1
424	6	2.2	165	14	US-10-106-698-6109	Sequence 6109, Ap	497	6	2.2	167	12	US-10-282-122A-62656	Sequence 62656, A
425	6	2.2	166	9	US-09-559-671A-75	Sequence 75, App1	498	6	2.2	167	12	US-10-424-599-211925	Sequence 211925, A
426	6	2.2	166	9	US-09-559-671A-76	Sequence 76, App1	499	6	2.2	167	12	US-10-424-599-211927	Sequence 211927, A
427	6	2.2	166	9	US-09-559-671A-77	Sequence 77, App1	500	6	2.2	167	15	US-10-369-493-1845	Sequence 1845, Ap
428	6	2.2	166	9	US-09-559-671A-79	Sequence 79, App1	501	6	2.2	169	12	US-10-282-122A-57648	Sequence 57648, A
429	6	2.2	166	9	US-09-559-671A-81	Sequence 81, App1	502	6	2.2	169	12	US-10-424-599-232864	Sequence 232864, A
430	6	2.2	166	9	US-09-559-671A-82	Sequence 82, App1	503	6	2.2	169	15	US-10-424-599-232864	Sequence 674, App
431	6	2.2	166	9	US-09-559-671A-84	Sequence 84, App1	504	6	2.2	170	9	US-09-989-919-78	Sequence 919, App1
432	6	2.2	166	9	US-09-559-671A-85	Sequence 85, App1	505	6	2.2	170	12	US-10-424-599-189920	Sequence 189920, A
433	6	2.2	166	9	US-09-559-671A-86	Sequence 86, App1	506	6	2.2	171	12	US-10-424-599-215472	Sequence 215472, A
434	6	2.2	166	9	US-09-977-034-2	Sequence 2, App1	507	6	2.2	171	12	US-10-424-599-177986	Sequence 177986, A
435	6	2.2	166	9	US-09-977-034-10	Sequence 10, App1	508	6	2.2	172	12	US-10-424-599-177986	Sequence 177986, A
436	6	2.2	166	9	US-09-977-034-13	Sequence 13, App1	509	6	2.2	172	15	US-10-369-493-15378	Sequence 15378, A
437	6	2.2	166	9	US-09-977-034-14	Sequence 14, App1	510	6	2.2	172	15	US-10-369-493-15378	Sequence 15378, A
438	6	2.2	166	9	US-09-977-034-15	Sequence 15, App1	511	6	2.2	173	12	US-10-369-493-15746	Sequence 15746, A
439	6	2.2	166	9	US-09-977-034-17	Sequence 17, App1	512	6	2.2	173	12	US-10-424-599-225389	Sequence 225389, A
440	6	2.2	166	9	US-09-977-034-18	Sequence 18, App1	513	6	2.2	174	12	US-10-424-599-176852	Sequence 176852, A
441	6	2.2	166	9	US-09-977-034-19	Sequence 19, App1	514	6	2.2	175	12	US-10-424-599-176852	Sequence 176852, A
442	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	515	6	2.2	176	14	US-10-127-032-122	Sequence 122, App
443	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	516	6	2.2	176	14	US-10-424-599-144059	Sequence 144059, A
444	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	517	6	2.2	177	12	US-10-424-599-144059	Sequence 144059, A
445	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	518	6	2.2	179	12	US-10-424-599-144059	Sequence 144059, A
446	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	519	6	2.2	179	12	US-10-424-599-144059	Sequence 144059, A
447	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	520	6	2.2	180	12	US-10-424-599-144059	Sequence 144059, A
448	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	521	6	2.2	180	12	US-10-424-599-144059	Sequence 144059, A
449	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	522	6	2.2	181	12	US-10-424-599-144059	Sequence 144059, A
450	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	523	6	2.2	182	9	US-09-791-171-12	Sequence 171, App1
451	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	524	6	2.2	182	9	US-09-791-171-12	Sequence 171, App1
452	6	2.2	166	10	US-09-800-505-109	Sequence 109, App	525	6	2.2	182	10	US-09-804-980-12	Sequence 12, App1
453	6	2.2	166	12	US-10-425-114-61680	Sequence 61680, A	526	6	2.2	182	12	US-10-424-599-185594	Sequence 185594, A

527	6	2.2	182	12	US-09-943-443-12	Sequence 12, Appl	600	6	2.2	210	12	US-10-435-114-51533	Sequence 51533, A
528	6	2.2	183	12	US-10-424-599-261099	Sequence 261099, A	601	6	2.2	212	12	US-10-282-122A-59410	Sequence 59410, A
529	6	2.2	183	12	US-10-424-599-285416	Sequence 285416, A	602	6	2.2	213	12	US-10-424-599-335392	Sequence 235392, A
530	6	2.2	185	12	US-10-425-114-45011	Sequence 45011, A	603	6	2.2	213	12	US-10-282-122A-44412	Sequence 44312, A
531	6	2.2	185	12	US-10-289-762-106	Sequence 106, Appl	604	6	2.2	213	12	US-10-424-599-145495	Sequence 145495, A
532	6	2.2	186	9	US-09-738-626-6231	Sequence 6231, Ap	605	6	2.2	213	12	US-10-291-190-48	Sequence 48, Appl
533	6	2.2	186	12	US-10-424-599-219603	Sequence 219603, A	606	6	2.2	214	14	US-10-032-585-7363	Sequence 763, Ap
534	6	2.2	186	12	US-10-158-057-223	Sequence 223, App	607	6	2.2	215	12	US-10-424-599-205345	Sequence 205345, A
535	6	2.2	187	9	US-09-925-937-490	Sequence 490, App	608	6	2.2	216	12	US-10-282-122A-51000	Sequence 51000, A
536	6	2.2	188	12	US-10-424-599-215561	Sequence 215561, A	609	6	2.2	216	12	US-10-424-599-175265	Sequence 8153, Ap
537	6	2.2	188	12	US-10-425-114-69592	Sequence 69592, A	610	6	2.2	216	14	US-10-128-714-8153	Sequence 8153, Ap
538	6	2.2	189	9	US-09-919-497-73	Sequence 73, Appl	611	6	2.2	217	14	US-10-376-345-4	Sequence 4, Appl
539	6	2.2	189	10	US-09-962-625-52	Sequence 2, Appl	612	6	2.2	217	14	US-10-376-345-6	Sequence 6, Appl
540	6	2.2	189	12	US-10-424-599-189299	Sequence 189299, A	613	6	2.2	218	12	US-10-425-114-45449	Sequence 44449, A
541	6	2.2	189	12	US-09-881-050-18	Sequence 18, Appl	614	6	2.2	220	13	US-09-726-643-54	Sequence 54, Appl
542	6	2.2	189	12	US-09-881-050-19	Sequence 19, Appl	615	6	2.2	220	13	US-10-042-141-54	Sequence 54, Appl
543	6	2.2	189	12	US-09-881-050-22	Sequence 22, Appl	616	6	2.2	221	10	US-09-934-455-216	Sequence 216, App
544	6	2.2	189	12	US-09-881-050-23	Sequence 23, Appl	617	6	2.2	221	12	US-10-282-122A-63302	Sequence 63302, A
545	6	2.2	189	12	US-09-881-050-25	Sequence 25, Appl	618	6	2.2	221	12	US-10-412-699B-1862	Sequence 1862, App
546	6	2.2	189	12	US-09-881-050-28	Sequence 28, Appl	619	6	2.2	221	12	US-10-235-066A-998	Sequence 998, App
547	6	2.2	189	14	US-10-284-740-10	Sequence 10, Appl	620	6	2.2	221	15	US-10-369-493-13371	Sequence 13371, A
548	6	2.2	189	16	US-10-415-963-38	Sequence 38, Appl	621	6	2.2	221	15	US-10-374-780A-274	Sequence 274, App
549	6	2.2	189	16	US-10-415-963-42	Sequence 42, Appl	622	6	2.2	223	12	US-10-424-599-147557	Sequence 147557, A
550	6	2.2	189	16	US-10-415-963-44	Sequence 44, Appl	623	6	2.2	223	12	US-10-276-774-2617	Sequence 2617, Ap
551	6	2.2	189	16	US-10-415-963-46	Sequence 46, Appl	624	6	2.2	223	14	US-10-156-761-14920	Sequence 14920, A
552	6	2.2	189	16	US-10-415-963-50	Sequence 50, Appl	625	6	2.2	224	9	US-09-393-634-9	Sequence 9, Appl
553	6	2.2	189	16	US-10-415-963-52	Sequence 52, Appl	626	6	2.2	224	12	US-10-364-861-9	Sequence 9, Appl
554	6	2.2	189	16	US-10-415-963-54	Sequence 54, Appl	627	6	2.2	224	14	US-10-383-982-9	Sequence 9, Appl
555	6	2.2	189	16	US-10-415-963-56	Sequence 56, Appl	628	6	2.2	226	9	US-09-815-242-12455	Sequence 12455, A
556	6	2.2	189	16	US-10-415-963-62	Sequence 62, Appl	629	6	2.2	226	9	US-09-815-242-12783	Sequence 12783, A
557	6	2.2	189	16	US-10-415-963-64	Sequence 64, Appl	630	6	2.2	226	15	US-10-369-493-8950	Sequence 8950, Ap
558	6	2.2	189	16	US-10-415-963-66	Sequence 66, Appl	631	6	2.2	227	15	US-10-346-000A-16	Sequence 16, Appl
559	6	2.2	189	16	US-10-415-963-68	Sequence 68, Appl	632	6	2.2	228	12	US-10-282-122A-50077	Sequence 50077, A
560	6	2.2	189	16	US-10-415-963-70	Sequence 70, Appl	633	6	2.2	229	12	US-10-335-977-7858	Sequence 7858, Ap
561	6	2.2	189	16	US-10-415-963-72	Sequence 72, Appl	634	6	2.2	229	15	US-10-369-493-6851	Sequence 6851, Ap
562	6	2.2	190	9	US-09-764-864-1071	Sequence 1071, Ap	635	6	2.2	230	9	US-09-738-626-5703	Sequence 5703, Ap
563	6	2.2	190	12	US-10-425-114-50711	Sequence 50711, A	636	6	2.2	231	14	US-10-106-698-6212	Sequence 6212, Ap
564	6	2.2	190	12	US-10-425-114-54773	Sequence 54773, A	637	6	2.2	231	14	US-10-107-857-48	Sequence 48, Appl
565	6	2.2	193	14	US-10-156-761-11378	Sequence 11378, A	638	6	2.2	231	16	US-10-389-566-933	Sequence 933, App
566	6	2.2	193	15	US-10-369-493-22176	Sequence 22176, A	639	6	2.2	232	12	US-10-424-599-145431	Sequence 145431, A
567	6	2.2	195	9	US-09-864-761-47064	Sequence 47064, A	640	6	2.2	234	12	US-10-282-122A-74300	Sequence 74300, A
568	6	2.2	196	12	US-10-424-599-175546	Sequence 175546, A	641	6	2.2	234	12	US-10-424-599-146518	Sequence 146518, A
569	6	2.2	196	12	US-10-376-774-2551	Sequence 2551, Ap	642	6	2.2	235	12	US-10-282-122A-67332	Sequence 67332, A
570	6	2.2	197	12	US-10-282-122A-71273	Sequence 71273, A	643	6	2.2	235	12	US-10-425-114-59338	Sequence 59338, A
571	6	2.2	198	9	US-09-925-302-500	Sequence 500, App	644	6	2.2	236	9	US-09-738-626-6304	Sequence 6304, Ap
572	6	2.2	198	12	US-09-925-302-500	Sequence 500, App	645	6	2.2	236	12	US-10-424-599-250081	Sequence 250081, A
573	6	2.2	199	15	US-10-369-493-17684	Sequence 17684, A	646	6	2.2	237	12	US-10-282-122A-77310	Sequence 77310, A
574	6	2.2	200	9	US-09-815-242-11021	Sequence 11021, A	647	6	2.2	237	12	US-10-425-114-41334	Sequence 41384, A
575	6	2.2	200	9	US-09-738-626-6458	Sequence 6458, Ap	648	6	2.2	237	12	US-10-425-114-56334	Sequence 56374, A
576	6	2.2	200	12	US-10-282-122A-52322	Sequence 52322, A	649	6	2.2	238	15	US-10-104-047-3302	Sequence 3302, Ap
577	6	2.2	200	12	US-10-282-122A-58168	Sequence 58168, A	650	6	2.2	239	12	US-10-424-599-229052	Sequence 229052, A
578	6	2.2	200	14	US-10-156-761-12379	Sequence 12379, A	651	6	2.2	239	12	US-10-424-599-282640	Sequence 282640, A
579	6	2.2	201	12	US-10-282-122A-62266	Sequence 62266, A	652	6	2.2	239	14	US-10-219-220-65	Sequence 65, Appl
580	6	2.2	201	14	US-10-128-714-3153	Sequence 3153, Ap	653	6	2.2	239	14	US-10-156-761-8154	Sequence 8154, Ap
581	6	2.2	202	12	US-10-335-977-8986	Sequence 8986, Ap	654	6	2.2	239	15	US-10-393-840-54	Sequence 854, Appl
582	6	2.2	203	8	US-08-781-986A-5207	Sequence 5207, Ap	655	6	2.2	239	15	US-10-393-840-117	Sequence 117, App
583	6	2.2	203	12	US-10-329-624-5207	Sequence 5207, Ap	656	6	2.2	240	14	US-10-369-493-23445	Sequence 23495, A
584	6	2.2	204	12	US-10-424-599-144986	Sequence 144986, A	657	6	2.2	240	15	US-10-156-761-8902	Sequence 8902, Ap
585	6	2.2	204	12	US-10-424-599-150135	Sequence 150135, A	658	6	2.2	243	9	US-09-738-626-4852	Sequence 4852, Ap
586	6	2.2	205	12	US-10-282-122A-67632	Sequence 67632, A	659	6	2.2	243	9	US-09-738-626-6413	Sequence 6413, Ap
587	6	2.2	205	12	US-10-424-599-248233	Sequence 248233, A	660	6	2.2	243	12	US-10-424-599-261883	Sequence 261883, A
588	6	2.2	205	12	US-10-425-114-63848	Sequence 63848, A	661	6	2.2	243	12	US-10-67-476-368	Sequence 368, App
589	6	2.2	207	12	US-10-282-122A-45062	Sequence 45062, A	662	6	2.2	243	12	US-10-344-980-2	Sequence 2, Appl
590	6	2.2	208	9	US-09-738-626-6261	Sequence 6261, Ap	663	6	2.2	244	15	US-10-369-493-16542	Sequence 16542, A
591	6	2.2	208	12	US-10-424-599-193167	Sequence 193167, A	664	6	2.2	245	10	US-09-903-463-15	Sequence 15, Appl
592	6	2.2	209	9	US-09-815-242-5583	Sequence 5583, Ap	665	6	2.2	245	12	US-10-425-114-43245	Sequence 43245, A
593	6	2.2	209	9	US-09-815-242-11967	Sequence 11967, A	666	6	2.2	246	12	US-10-282-122A-44007	Sequence 44007, A
594	6	2.2	209	9	US-09-738-626-5728	Sequence 5728, Ap	667	6	2.2	247	12	US-10-424-599-155063	Sequence 155063, A
595	6	2.2	209	12	US-10-282-122A-66579	Sequence 66579, A	668	6	2.2	248	12	US-10-282-122A-57202	Sequence 57202, A
596	6	2.2	209	12	US-10-424-599-23311	Sequence 23311, A	669	6	2.2	248	14	US-10-156-761-11171	Sequence 11171, A
597	6	2.2	209	14	US-10-307-294-7	Sequence 7, Appl	670	6	2.2	249	15	US-10-369-493-6091	Sequence 6091, Ap
598	6	2.2	210	10	US-09-975-719-172	Sequence 172, App	671	6	2.2	249	15	US-09-984-186-18	Sequence 18, Appl
599	6	2.2	210	12	US-10-282-122A-55523	Sequence 55523, A	672	6	2.2	249	14	US-10-217-667-18	Sequence 18, Appl

673	6	2.2	249	14	US-10-237-708-18	Sequence 18, Appl	746	6	2.2	277	12	US-10-206-576-30	Sequence 30, Appl
674	6	2.2	249	14	US-10-237-866-18	Sequence 18, Appl	747	6	2.2	278	15	US-10-369-493-8275	Sequence 8275, Ap
675	6	2.2	249	14	US-10-237-871-18	Sequence 18, Appl	748	6	2.2	279	12	US-10-282-1228-63886	Sequence 63886, A
676	6	2.2	249	14	US-10-237-624-18	Sequence 18, Appl	749	6	2.2	279	14	US-10-156-761-14320	Sequence 14320, A
677	6	2.2	249	14	US-10-156-761-12944	Sequence 12944, A	750	6	2.2	280	8	US-08-781-986A-5195	Sequence 5195, Ap
678	6	2.2	249	16	US-10-702-536-18	Sequence 18, Appl	751	6	2.2	280	12	US-10-424-599-252283	Sequence 252283,
679	6	2.2	249	16	US-10-702-636-18	Sequence 18, Appl	752	6	2.2	280	12	US-10-339-624-5195	Sequence 5195, Ap
680	6	2.2	250	10	US-09-769-736-98	Sequence 98, Appl	753	6	2.2	280	16	US-10-408-765A-151	Sequence 151, Appl
681	6	2.2	250	12	US-10-424-599-278915	Sequence 178915,	754	6	2.2	281	14	US-10-376-345-2	Sequence 2, Appl
682	6	2.2	250	12	US-10-424-599-284325	Sequence 284325,	755	6	2.2	281	10	US-09-769-736-54	Sequence 54, Appl
683	6	2.2	251	14	US-10-189-346-86	Sequence 86, Appl	756	6	2.2	282	16	US-10-389-566-1240	Sequence 1240, Ap
684	6	2.2	251	15	US-10-310-154-707	Sequence 707, App	757	6	2.2	283	14	US-10-156-761-13731	Sequence 12731, A
685	6	2.2	252	15	US-10-282-1228-70252	Sequence 70252, A	758	6	2.2	284	9	US-09-815-242-13173	Sequence 11773, A
686	6	2.2	252	12	US-10-424-599-180645	Sequence 180645,	759	6	2.2	284	12	US-10-282-1228-73708	Sequence 73708, A
687	6	2.2	253	12	US-10-424-599-272981	Sequence 272981,	760	6	2.2	284	12	US-10-627-176-446	Sequence 446, App
688	6	2.2	254	12	US-10-282-1228-69050	Sequence 69050, A	761	6	2.2	284	12	US-10-627-176-446	Sequence 450, App
689	6	2.2	254	12	US-10-425-114-68110	Sequence 68110, A	762	6	2.2	284	14	US-10-156-761-10979	Sequence 10979, A
690	6	2.2	255	15	US-10-369-493-21110	Sequence 21110, A	763	6	2.2	284	15	US-10-369-493-7492	Sequence 4731, Ap
691	6	2.2	257	12	US-10-206-576-32	Sequence 32, Appl	764	6	2.2	285	15	US-10-369-493-8027	Sequence 8027, Ap
692	6	2.2	258	9	US-09-864-761-37821	Sequence 37821, A	765	6	2.2	285	15	US-10-369-493-8027	Sequence 450, App
693	6	2.2	258	12	US-10-425-114-52260	Sequence 52260, A	766	6	2.2	285	15	US-10-369-493-13392	Sequence 13392, A
694	6	2.2	258	12	US-10-425-114-62204	Sequence 62204, A	767	6	2.2	285	16	US-10-432-443-146	Sequence 146, App
695	6	2.2	259	9	US-09-815-242-5146	Sequence 5146, Ap	768	6	2.2	286	12	US-10-282-1228-62914	Sequence 62914, A
696	6	2.2	259	12	US-10-282-1228-43469	Sequence 43469, A	769	6	2.2	287	12	US-10-282-1228-65725	Sequence 65725, A
697	6	2.2	259	12	US-10-282-1228-63628	Sequence 63628, A	770	6	2.2	287	12	US-10-282-1228-78036	Sequence 78036, A
698	6	2.2	259	16	US-10-432-443-133	Sequence 133, App	771	6	2.2	288	12	US-10-282-1228-49435	Sequence 49435, A
699	6	2.2	260	9	US-09-765-772-32	Sequence 32, Appl	772	6	2.2	288	14	US-10-043-487-366	Sequence 366, App
700	6	2.2	260	12	US-10-282-1228-45357	Sequence 45357, A	773	6	2.2	288	15	US-10-369-493-9706	Sequence 9706, App
701	6	2.2	260	12	US-10-282-1228-69360	Sequence 69360, A	774	6	2.2	288	15	US-10-425-114-40612	Sequence 40612, A
702	6	2.2	260	12	US-10-282-1228-70210	Sequence 70210, A	775	6	2.2	289	12	US-10-425-114-40612	Sequence 13759, A
703	6	2.2	260	12	US-10-424-599-274453	Sequence 274453, A	776	6	2.2	289	14	US-10-156-761-13759	Sequence 13759, A
704	6	2.2	261	12	US-10-282-1228-67692	Sequence 67692, A	777	6	2.2	289	15	US-10-369-493-18920	Sequence 18920, A
705	6	2.2	261	14	US-10-161-051-23	Sequence 23, Appl	778	6	2.2	290	12	US-10-282-1228-53686	Sequence 53686, A
706	6	2.2	261	15	US-10-369-493-7775	Sequence 7775, Ap	779	6	2.2	291	12	US-10-424-599-204055	Sequence 204055,
707	6	2.2	261	15	US-10-264-213-227	Sequence 227, App	780	6	2.2	291	12	US-10-424-599-207605	Sequence 207605,
708	6	2.2	262	12	US-10-282-1228-54537	Sequence 54537, A	781	6	2.2	292	12	US-10-156-761-14678	Sequence 14678, A
709	6	2.2	262	14	US-10-156-761-10500	Sequence 10500, A	782	6	2.2	292	14	US-10-335-977-5290	Sequence 5290, Ap
710	6	2.2	263	12	US-10-282-1228-48062	Sequence 48062, A	783	6	2.2	293	15	US-10-369-493-23620	Sequence 23620, A
711	6	2.2	264	12	US-10-282-1228-71721	Sequence 71721, A	784	6	2.2	294	16	US-10-389-566-1246	Sequence 1246, Ap
712	6	2.2	265	12	US-10-424-599-258008	Sequence 258008,	785	6	2.2	294	16	US-10-389-566-1606	Sequence 1606, Ap
713	6	2.2	265	12	US-10-335-977-5161	Sequence 5161, Ap	786	6	2.2	295	16	US-10-389-566-645	Sequence 645, App
714	6	2.2	265	12	US-10-335-977-5332	Sequence 5332, Ap	787	6	2.2	296	9	US-09-815-242-1246	Sequence 5446, App
715	6	2.2	266	12	US-10-335-977-5332	Sequence 5332, Ap	788	6	2.2	296	9	US-09-815-242-1246	Sequence 6445, App
716	6	2.2	266	12	US-10-335-977-5160	Sequence 51604, Ap	789	6	2.2	296	14	US-10-029-186-32377	Sequence 32377, A
717	6	2.2	266	15	US-10-156-761-15054	Sequence 15054, A	790	6	2.2	296	14	US-10-025-806-116	Sequence 116, App
718	6	2.2	266	15	US-10-320-797-3067	Sequence 3067, Ap	791	6	2.2	296	14	US-10-374-780A-824	Sequence 824, App
719	6	2.2	267	12	US-10-282-1228-53361	Sequence 53361, A	792	6	2.2	297	12	US-10-425-114-58061	Sequence 58061, A
720	6	2.2	267	12	US-10-424-599-273056	Sequence 273056,	793	6	2.2	297	12	US-10-369-493-4306	Sequence 4306, Ap
721	6	2.2	268	9	US-09-815-242-11314	Sequence 11314, A	794	6	2.2	298	15	US-10-282-1228-57061	Sequence 57061, A
722	6	2.2	268	10	US-09-815-242-11495	Sequence 11495, A	795	6	2.2	298	12	US-10-424-599-187636	Sequence 187636,
723	6	2.2	268	12	US-10-282-1228-58680	Sequence 58680, A	796	6	2.2	298	12	US-09-815-242-4	Sequence 4, Appl
724	6	2.2	268	12	US-10-335-977-5339	Sequence 5339, Ap	797	6	2.2	298	15	US-10-369-493-6885	Sequence 6885, Ap
725	6	2.2	268	12	US-10-282-1228-46819	Sequence 46819, A	798	6	2.2	298	16	US-10-389-566-550	Sequence 550, App
726	6	2.2	270	12	US-10-282-1228-70899	Sequence 70899, A	799	6	2.2	298	16	US-10-389-566-644	Sequence 644, App
727	6	2.2	270	12	US-10-282-1228-71810	Sequence 71810, A	800	6	2.2	299	9	US-09-738-626-1209	Sequence 4209, Ap
728	6	2.2	271	12	US-10-335-977-5334	Sequence 5334, Ap	801	6	2.2	300	12	US-10-425-114-59605	Sequence 59605, A
729	6	2.2	271	12	US-10-282-1228-50375	Sequence 50375, A	802	6	2.2	300	15	US-10-369-493-23126	Sequence 23126, A
730	6	2.2	272	12	US-10-282-1228-51605	Sequence 51605, A	803	6	2.2	300	15	US-10-369-493-23126	Sequence 23126, A
731	6	2.2	272	15	US-10-369-493-12081	Sequence 12081, A	804	6	2.2	302	9	US-09-815-242-12336	Sequence 12336, A
732	6	2.2	273	15	US-10-282-1228-65474	Sequence 65474, A	805	6	2.2	302	9	US-09-815-242-12336	Sequence 12336, A
733	6	2.2	274	12	US-10-282-1228-65474	Sequence 65474, A	806	6	2.2	302	12	US-10-282-1228-44027	Sequence 44027, A
734	6	2.2	274	12	US-10-282-1228-65474	Sequence 65474, A	807	6	2.2	302	12	US-10-282-1228-70930	Sequence 70930, A
735	6	2.2	275	12	US-10-282-1228-66155	Sequence 66155, A	808	6	2.2	302	12	US-10-424-599-275275	Sequence 275275,
736	6	2.2	275	12	US-10-425-114-68522	Sequence 68522, A	809	6	2.2	302	12	US-10-424-599-275275	Sequence 16, Appl
737	6	2.2	275	14	US-10-295-220-26	Sequence 26, Appl	810	6	2.2	303	9	US-09-738-029-16	Sequence 16, Appl
738	6	2.2	275	15	US-10-369-493-13473	Sequence 13273, A	811	6	2.2	303	12	US-10-004-860-818	Sequence 818, App
739	6	2.2	276	15	US-10-264-213-242	Sequence 242, App	812	6	2.2	303	12	US-10-004-860-818	Sequence 818, App
740	6	2.2	277	8	US-08-834-705-24	Sequence 24, Appl	813	6	2.2	304	14	US-10-023-882-818	Sequence 818, App
741	6	2.2	277	12	US-10-282-1228-57209	Sequence 57209, A	814	6	2.2	304	14	US-10-425-114-59882	Sequence 59872, A
742	6	2.2	277	12	US-10-282-1228-58897	Sequence 58897, A	815	6	2.2	305	12	US-10-425-114-66119	Sequence 66119, A
743	6	2.2	277	12	US-10-424-599-160702	Sequence 160702,	816	6	2.2	305	15	US-10-369-493-162	Sequence 162, App
744	6	2.2	277	12	US-10-425-114-43859	Sequence 43859, A	817	6	2.2	305	15	US-10-369-493-11533	Sequence 11533, A
745	6	2.2	277	12	US-10-335-977-5289	Sequence 5289, Ap	818	6	2.2	305	15	US-10-369-493-14443	Sequence 14443, A

819	6	2.2	305	15	US-10-369-493-14764	Sequence 14764, A	892	6	2.2	307	14	US-10-242-990-30	Sequence 30, Appl
820	6	2.2	305	15	US-10-369-493-15244	Sequence 15244, A	893	6	2.2	307	14	US-10-243-023-30	Sequence 30, Appl
821	6	2.2	306	10	US-09-769-736-112	Sequence 112, Appl	894	6	2.2	307	14	US-10-243-023-30	Sequence 30, Appl
822	6	2.2	306	12	US-09-758-759-115	Sequence 115, Appl	895	6	2.2	307	14	US-10-243-026-30	Sequence 30, Appl
823	6	2.2	306	15	US-10-369-493-7062	Sequence 7062, Ap	896	6	2.2	307	14	US-10-243-326-30	Sequence 30, Appl
824	6	2.2	307	12	US-10-245-752-30	Sequence 74940, A	897	6	2.2	307	14	US-10-243-364-30	Sequence 30, Appl
825	6	2.2	307	12	US-10-245-752-30	Sequence 30, Appl	898	6	2.2	307	14	US-10-243-494-30	Sequence 30, Appl
826	6	2.2	307	12	US-10-245-859-30	Sequence 28, Appl	899	6	2.2	307	14	US-10-243-995-30	Sequence 30, Appl
827	6	2.2	307	13	US-10-001-054-28	Sequence 30, Appl	900	6	2.2	307	14	US-10-245-230-30	Sequence 30, Appl
828	6	2.2	307	14	US-10-245-103-30	Sequence 30, Appl	901	6	2.2	307	14	US-10-245-253-30	Sequence 30, Appl
829	6	2.2	307	14	US-10-245-107-30	Sequence 30, Appl	902	6	2.2	307	14	US-10-245-479-30	Sequence 30, Appl
830	6	2.2	307	14	US-10-245-143-30	Sequence 30, Appl	903	6	2.2	307	14	US-10-245-499-30	Sequence 30, Appl
831	6	2.2	307	14	US-10-245-771-30	Sequence 30, Appl	904	6	2.2	307	14	US-10-245-772-30	Sequence 30, Appl
832	6	2.2	307	14	US-10-245-851-30	Sequence 30, Appl	905	6	2.2	307	14	US-10-245-811-30	Sequence 30, Appl
833	6	2.2	307	14	US-10-245-883-30	Sequence 30, Appl	906	6	2.2	307	14	US-10-245-812-30	Sequence 30, Appl
834	6	2.2	307	14	US-10-237-553-30	Sequence 30, Appl	907	6	2.2	307	14	US-10-245-852-30	Sequence 30, Appl
835	6	2.2	307	14	US-10-238-183-30	Sequence 30, Appl	908	6	2.2	307	14	US-10-245-875-30	Sequence 30, Appl
836	6	2.2	307	14	US-10-238-283-30	Sequence 30, Appl	909	6	2.2	307	14	US-10-245-881-30	Sequence 30, Appl
837	6	2.2	307	14	US-10-238-370-30	Sequence 30, Appl	910	6	2.2	307	14	US-10-245-911-30	Sequence 30, Appl
838	6	2.2	307	14	US-10-245-055-30	Sequence 30, Appl	911	6	2.2	307	14	US-10-245-913-30	Sequence 30, Appl
839	6	2.2	307	14	US-10-245-147-30	Sequence 30, Appl	912	6	2.2	307	14	US-10-246-080-30	Sequence 30, Appl
840	6	2.2	307	14	US-10-245-730-30	Sequence 30, Appl	913	6	2.2	307	14	US-10-246-121-30	Sequence 30, Appl
841	6	2.2	307	14	US-10-245-739-30	Sequence 30, Appl	914	6	2.2	307	14	US-10-246-305-30	Sequence 30, Appl
842	6	2.2	307	14	US-10-246-210-30	Sequence 30, Appl	915	6	2.2	307	14	US-10-246-829-30	Sequence 30, Appl
843	6	2.2	307	14	US-10-239-196-30	Sequence 30, Appl	916	6	2.2	307	14	US-10-247-036-30	Sequence 30, Appl
844	6	2.2	307	14	US-10-243-024-30	Sequence 30, Appl	917	6	2.2	307	14	US-10-243-255-30	Sequence 30, Appl
845	6	2.2	307	14	US-10-243-409-30	Sequence 30, Appl	918	6	2.2	307	14	US-10-245-810-30	Sequence 30, Appl
846	6	2.2	307	14	US-10-245-621-30	Sequence 30, Appl	919	6	2.2	307	14	US-10-245-910-30	Sequence 30, Appl
847	6	2.2	307	14	US-10-245-880-30	Sequence 30, Appl	920	6	2.2	307	14	US-10-246-098-30	Sequence 30, Appl
848	6	2.2	307	14	US-10-245-031-30	Sequence 30, Appl	921	6	2.2	307	14	US-10-237-496-30	Sequence 30, Appl
849	6	2.2	307	14	US-10-243-093-30	Sequence 30, Appl	922	6	2.2	307	14	US-10-242-074-30	Sequence 30, Appl
850	6	2.2	307	14	US-10-243-183-30	Sequence 30, Appl	923	6	2.2	307	14	US-10-242-505-30	Sequence 30, Appl
851	6	2.2	307	14	US-10-245-427-30	Sequence 30, Appl	924	6	2.2	307	14	US-10-242-574-30	Sequence 30, Appl
852	6	2.2	307	14	US-10-245-473-30	Sequence 30, Appl	925	6	2.2	307	14	US-10-243-261-30	Sequence 30, Appl
853	6	2.2	307	14	US-10-245-770-30	Sequence 30, Appl	926	6	2.2	307	14	US-10-243-282-30	Sequence 30, Appl
854	6	2.2	307	14	US-10-245-877-30	Sequence 30, Appl	927	6	2.2	307	14	US-10-243-402-30	Sequence 30, Appl
855	6	2.2	307	14	US-10-246-976-30	Sequence 30, Appl	928	6	2.2	307	14	US-10-243-431-30	Sequence 30, Appl
856	6	2.2	307	14	US-10-243-329-30	Sequence 30, Appl	929	6	2.2	307	14	US-10-244-164-30	Sequence 30, Appl
857	6	2.2	307	14	US-10-242-743-30	Sequence 30, Appl	930	6	2.2	307	14	US-10-244-972-30	Sequence 30, Appl
858	6	2.2	307	14	US-10-242-845-30	Sequence 30, Appl	931	6	2.2	307	14	US-10-197-942-30	Sequence 30, Appl
859	6	2.2	307	14	US-10-237-636-30	Sequence 30, Appl	932	6	2.2	307	14	US-10-238-196-30	Sequence 30, Appl
860	6	2.2	307	14	US-10-238-325-30	Sequence 30, Appl	933	6	2.2	307	14	US-10-245-013-30	Sequence 30, Appl
861	6	2.2	307	14	US-10-238-346-30	Sequence 30, Appl	934	6	2.2	307	15	US-10-264-213-150	Sequence 150, App
862	6	2.2	307	14	US-10-238-411-30	Sequence 30, Appl	935	6	2.2	308	9	US-09-815-242-13087	Sequence 13087, A
863	6	2.2	307	14	US-10-243-128-30	Sequence 30, Appl	936	6	2.2	308	12	US-10-282-1228-72664	Sequence 72664, A
864	6	2.2	307	14	US-10-243-425-30	Sequence 30, Appl	937	6	2.2	308	12	US-10-107-431-25	Sequence 25, Appl
865	6	2.2	307	14	US-10-243-446-30	Sequence 30, Appl	938	6	2.2	308	12	US-09-758-759-11	Sequence 11, Appl
866	6	2.2	307	14	US-10-245-874-30	Sequence 30, Appl	939	6	2.2	308	14	US-10-156-761-11436	Sequence 11436, A
867	6	2.2	307	14	US-10-242-653-30	Sequence 30, Appl	940	6	2.2	309	9	US-09-815-242-13711	Sequence 13711, A
868	6	2.2	307	14	US-10-243-167-30	Sequence 30, Appl	941	6	2.2	309	10	US-09-510-332-85	Sequence 85, Appl
869	6	2.2	307	14	US-10-243-389-30	Sequence 30, Appl	942	6	2.2	309	10	US-09-510-332-131	Sequence 131, App
870	6	2.2	307	14	US-10-244-947-30	Sequence 30, Appl	943	6	2.2	309	12	US-10-282-1228-45051	Sequence 45051, A
871	6	2.2	307	14	US-10-244-966-30	Sequence 30, Appl	944	6	2.2	309	12	US-10-282-1228-64247	Sequence 64247, A
872	6	2.2	307	14	US-10-244-079-30	Sequence 30, Appl	945	6	2.2	309	12	US-10-282-1228-74839	Sequence 74839, A
873	6	2.2	307	14	US-10-245-079-30	Sequence 30, Appl	946	6	2.2	309	12	US-10-282-1228-75392	Sequence 75392, A
874	6	2.2	307	14	US-10-245-127-30	Sequence 30, Appl	947	6	2.2	309	14	US-10-309-933-3	Sequence 3, Appl
875	6	2.2	307	14	US-10-245-207-30	Sequence 30, Appl	948	6	2.2	310	8	US-08-964-716-42	Sequence 42, Appl
876	6	2.2	307	14	US-10-246-646-30	Sequence 30, Appl	949	6	2.2	310	9	US-09-815-242-10667	Sequence 10667, A
877	6	2.2	307	14	US-10-246-695-30	Sequence 30, Appl	950	6	2.2	310	12	US-10-282-1228-72138	Sequence 72138, A
878	6	2.2	307	14	US-10-246-699-30	Sequence 30, Appl	951	6	2.2	310	15	US-10-369-493-1160	Sequence 1160, Ap
879	6	2.2	307	14	US-10-246-737-30	Sequence 30, Appl	952	6	2.2	311	12	US-10-282-1228-49082	Sequence 49082, A
880	6	2.2	307	14	US-10-246-878-30	Sequence 30, Appl	953	6	2.2	311	12	US-10-282-1228-63438	Sequence 63438, A
881	6	2.2	307	14	US-10-246-890-30	Sequence 30, Appl	954	6	2.2	311	12	US-10-424-599-261206	Sequence 261206, A
882	6	2.2	307	14	US-10-246-899-30	Sequence 30, Appl	955	6	2.2	311	14	US-10-156-761-10112	Sequence 10112, A
883	6	2.2	307	14	US-10-246-900-30	Sequence 30, Appl	956	6	2.2	312	12	US-10-282-1228-75790	Sequence 75790, A
884	6	2.2	307	14	US-10-247-098-30	Sequence 30, Appl	957	6	2.2	315	12	US-10-282-1228-52547	Sequence 52547, A
885	6	2.2	307	14	US-10-246-454-30	Sequence 30, Appl	958	6	2.2	315	12	US-10-425-114-52560	Sequence 52560, A
886	6	2.2	307	14	US-10-237-471-30	Sequence 30, Appl	959	6	2.2	315	14	US-10-032-585-7236	Sequence 7236, Ap
887	6	2.2	307	14	US-10-238-261-30	Sequence 30, Appl	960	6	2.2	316	12	US-10-282-1228-61028	Sequence 61028, A
888	6	2.2	307	14	US-10-238-324-30	Sequence 30, Appl	961	6	2.2	316	15	US-10-119-220-260	Sequence 260, App
889	6	2.2	307	14	US-10-241-860-30	Sequence 30, Appl	962	6	2.2	317	14	US-10-369-493-22263	Sequence 22263, A
890	6	2.2	307	14	US-10-242-172-30	Sequence 30, Appl	963	6	2.2	317	5	US-08-815-242-13654	Sequence 13654, A
891	6	2.2	307	14	US-10-242-654-30	Sequence 30, Appl	964	6	2.2	317	10	US-09-769-744A-116	Sequence 116, App

665 6 2.2 317 12 US-10-282-122A-72945 Sequence 72945, A
666 6 2.2 317 12 US-10-282-122A-74193 Sequence 74193, A
667 6 2.2 318 9 US-09-815-242-13749 Sequence 13749, A
668 6 2.2 318 10 US-09-815-242-13749 Sequence 90, Appl
669 6 2.2 318 12 US-10-282-122A-77438 Sequence 77438, A
670 6 2.2 318 12 US-10-225-066A-524 Sequence 524, Appl
671 6 2.2 318 15 US-10-374-780A-2328 Sequence 2328, Ap
672 6 2.2 320 12 US-10-424-599-188103 Sequence 188103, A
673 6 2.2 320 12 US-10-282-122A-49592 Sequence 49592, A
674 6 2.2 321 12 US-10-282-122A-59786 Sequence 59786, A
675 6 2.2 321 14 US-10-029-386-31960 Sequence 31960, A
676 6 2.2 321 15 US-10-369-493-2285 Sequence 2285, Ap
677 6 2.2 322 15 US-10-369-493-17239 Sequence 17239, Ap
678 6 2.2 322 15 US-10-374-780A-1703 Sequence 1703, Ap
679 6 2.2 323 9 US-09-886-055-339 Sequence 339, Appl
680 6 2.2 323 10 US-09-804-291-339 Sequence 339, Appl
681 6 2.2 323 12 US-10-282-122A-59321 Sequence 59321, A
682 6 2.2 323 12 US-10-282-122A-74657 Sequence 74657, A
683 6 2.2 323 12 US-10-424-599-196294 Sequence 196294, A
684 6 2.2 323 12 US-10-343-650A-136 Sequence 136, Appl
685 6 2.2 323 14 US-10-017-161-896 Sequence 896, Appl
686 6 2.2 323 15 US-10-024-212-32 Sequence 32, Appl
687 6 2.2 323 15 US-10-292-798-772 Sequence 772, Appl
688 6 2.2 324 12 US-10-282-122A-48691 Sequence 48691, A
689 6 2.2 324 14 US-10-156-761-9761 Sequence 9761, Ap
690 6 2.2 326 12 US-10-424-599-153681 Sequence 153681, A
691 6 2.2 326 15 US-10-369-493-6333 Sequence 6333, Ap
692 6 2.2 326 15 US-10-369-493-21352 Sequence 21352, A
693 6 2.2 327 12 US-10-424-599-239310 Sequence 239310, A
694 6 2.2 327 12 US-10-425-114-43573 Sequence 43573, A
695 6 2.2 327 12 US-10-425-114-43573 Sequence 46300, A
696 6 2.2 327 12 US-10-369-493-3791 Sequence 3791, Ap
697 6 2.2 328 12 US-10-282-122A-58389 Sequence 58389, A
698 6 2.2 328 12 US-10-425-114-52030 Sequence 52030, A
699 6 2.2 329 9 US-09-815-242-13786 Sequence 13786, A
1000 6 2.2 329 9 US-09-977-577-9 Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-10-282-122A-63386
Sequence 63386, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578

665 PRIOR FILING DATE: 2000-10-23
666 PRIOR APPLICATION NUMBER: 60/253,625
667 PRIOR FILING DATE: 2000-11-27
668 PRIOR APPLICATION NUMBER: 60/257,931
669 PRIOR FILING DATE: 2000-12-22
670 PRIOR APPLICATION NUMBER: 60/267,636
671 PRIOR FILING DATE: 2001-02-09
672 PRIOR APPLICATION NUMBER: 60/269,308
673 PRIOR FILING DATE: 2001-02-16
674 Remaining Prior Application data removed - See File Wrapper or PALM.
675 NUMBER OF SEQ ID NOS: 78614
676 SOFTWARE: PatentIn version 3.1
677 SEQ ID NO: 63386
678 LENGTH: 276
679 TYPE: PRT
680 ORGANISM: Moraxella catarrhalis
681 US-10-282-122A-63386
Query Match 88.0%; Score 243; DB 12; Length 276;
Best Local Similarity 100.0%; Pred. No. 1,7e-227;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
34 TAAQTIRGVAGAGPQAVAEVAGVAKRNLTVELVEFNDYAMPNSAVSKGEIDANAMQ 93
34 TAAQTIRGVAGAGPQAVAEVAGVAKRNLTVELVEFNDYAMPNSAVSKGEIDANAMQ 93
94 HKPYEKDSQSKANNLVYGNTPYPLAGYSTKIKTINELKDGATTAVPDPSNLARAL 153
94 HKPYEKDSQSKANNLVYGNTPYPLAGYSTKIKTINELKDGATTAVPDPSNLARAL 153
154 ILLEKQGIKIKDNTNLFSTLIDIVENPKKLVKEVDTSVAAAIIDVDIAVNNNTAQ 213
154 ILLEKQGIKIKDNTNLFSTLIDIVENPKKLVKEVDTSVAAAIIDVDIAVNNNTAQ 213
214 VGLTSENSEGVPEKDSRYNIIYARADNKSRAIOPEVAYCQDEVEAEKQFQKQVI 273
214 VGLTSENSEGVPEKDSRYNIIYARADNKSRAIOPEVAYCQDEVEAEKQFQKQVI 273
QY 274 KGM 276
DB 274 KGM 276
274 KGM 276
274 KGM 276
RESULT 2
US-09-815-242-11067
Sequence 11067, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: Prokaryotes
CURRENT APPLICATION NUMBER: US/09/815.242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22


```
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 1410
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11067
;; LENGTH: 273
;; TYPE: PRT
;; ORGANISM: Haemophilus influenzae
US-09-815-242-11067
```

```
Query Match          6.9%; Score 19; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      107 LNNLVIVGNTFYVPLAGYS 125
Db      104 LNNLVIVGNTFYVPLAGYS 122
```

```
RESULT 3
US-10-282-122A-58240
; Sequence 58240, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
```

```
;; APPLICANT: Wang, Liangsu
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Malone, Cheryl
;; APPLICANT: Haseibeck, Robert
;; APPLICANT: Ohlsen, Karl
;; APPLICANT: Zyskind, Judith
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John
;; APPLICANT: Carr, Grant
;; APPLICANT: Yamamoto, Robert
;; APPLICANT: Forsyth, R.
;; APPLICANT: Xu, H.
;; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;; FILE REFERENCE: EITPA.034A
;; CURRENT APPLICATION NUMBER: US/10/282,122A
;; CURRENT FILING DATE: 2003-02-20
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 58240
;; LENGTH: 273
;; TYPE: PRT
;; ORGANISM: Haemophilus influenzae
US-10-282-122A-58240
```

```
Query Match          6.9%; Score 19; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      107 LNNLVIVGNTFYVPLAGYS 125
Db      104 LNNLVIVGNTFYVPLAGYS 122
```

```
RESULT 4
US-10-380-817-2
; Sequence 2, Application US/10380817
; Publication No. US20040039169A1
; GENERAL INFORMATION:
;; APPLICANT: Thonard, Joelle
;; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
;; FILE REFERENCE: BM45419
;; CURRENT APPLICATION NUMBER: US/10/380,817
;; CURRENT FILING DATE: 2003-03-18
;; PRIOR APPLICATION NUMBER: PCT/EP01/10979
;; PRIOR FILING DATE: 2001-09-18
;; PRIOR APPLICATION NUMBER: GB 0022992.2
;; PRIOR FILING DATE: 2000-09-19
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 273
;; TYPE: PRT
;; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-2
```

```
Query Match          6.9%; Score 19; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      107 LNNLVIVGNTFYVPLAGYS 125
Db      104 LNNLVIVGNTFYVPLAGYS 122
```

```
RESULT 5
US-10-380-817-4
; Sequence 4, Application US/10380817
; Publication No. US20040039169A1
; GENERAL INFORMATION:
;; APPLICANT: Thonard, Joelle
;; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
;; FILE REFERENCE: BM45419
;; CURRENT APPLICATION NUMBER: US/10/380,817
;; CURRENT FILING DATE: 2003-03-18
;; PRIOR APPLICATION NUMBER: PCT/EP01/10979
;; PRIOR FILING DATE: 2001-09-18
;; PRIOR APPLICATION NUMBER: GB 0022992.2
;; PRIOR FILING DATE: 2000-09-19
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 273
;; TYPE: PRT
;; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-4
```

```
Query Match          6.9%; Score 19; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      107 LNNLVIVGNTFYVPLAGYS 125
Db      104 LNNLVIVGNTFYVPLAGYS 122
```

```
RESULT 6
US-10-380-817-6
; Sequence 6, Application US/10380817
; Publication No. US20040039169A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
; TITLE OF INVENTION: POLYPEPTIDE, PRODUCTION, VACCINE AND DIAGNOSTIC USE
; FILE REFERENCE: BM45419
; CURRENT APPLICATION NUMBER: US/10/380,817
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: PCT/EP01/10979
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: GB 0022992.2
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 273
; TYPE: PRT
; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-6

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 273;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFVYPLAGYS 125
Db 104 LNNLVIVGNTFVYPLAGYS 122

RESULT 7
US-10-380-817-8
; Sequence 8, Application US/10380817
; Publication No. US20040039169A1
; GENERAL INFORMATION:
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
; FILE REFERENCE: BM45419
; CURRENT APPLICATION NUMBER: US/10/380,817
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: PCT/EP01/10979
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: GB 0022992.2
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 273
; TYPE: PRT
; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-8

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 273;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFVYPLAGYS 125
Db 104 LNNLVIVGNTFVYPLAGYS 122

RESULT 8
US-10-380-817-10
; Sequence 10, Application US/10380817
; Publication No. US20040039169A1
; GENERAL INFORMATION:
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: HAEMOPHILUS INFLUENZAE BASB202
; FILE REFERENCE: BM45419
; CURRENT APPLICATION NUMBER: US/10/380,817
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: PCT/EP01/10979
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/EP01/10979
; PRIOR FILING DATE: 2001-09-18
```

```

; PRIOR APPLICATION NUMBER: GB 0022992.2
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10
; LENGTH: 273
; TYPE: PRT
; ORGANISM: No. US20040039169A1-typeable Haemophilus influenzae
US-10-380-817-10

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 273;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIVGNTFVYPLAGYS 125
Db 104 LNNLVIVGNTFVYPLAGYS 122

RESULT 9
US-10-282-122A-68391
; Sequence 68391, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27/257,931
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 68391
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-282-122A-68391

Query Match
Best Local Similarity 100.0%; Score 16; DB 12; Length 271;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 224 FVEKDSYNNIIVAR 239
DB 219 FVEKDSYNNIIVAR 234

RESULT 10

US-10-282-122A-52838
Sequence 52838, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52838
LENGTH: 270
TYPE: PRT
ORGANISM: Clostridium botulinum
US-10-282-122A-52838

Query Match 4.7%; Score 13; DB 12; Length 270;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 LKDGATIAVNDP 146
DB 130 LKDGATIAVNDP 142

RESULT 11

US-10-282-122A-67387
Sequence 67387, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 67387
LENGTH: 276
TYPE: PRT
ORGANISM: Pasteurella multocida
US-10-282-122A-67387

Query Match 4.3%; Score 12; DB 12; Length 276;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LLEKGLIKKD 166
DB 155 LLEKGLIKKD 166

RESULT 12

US-10-092-243A-15
Sequence 15, Application US/10092243A
Publication No. US20020197625A1
GENERAL INFORMATION:
APPLICANT: Hallman, Jeffrey D.
TITLE OF INVENTION: Microbial Polynucleotides Expressed During Infection of
FILE REFERENCE: MBH00-5051
CURRENT FILING DATE: 2002-05-21
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: 60/147,551
PRIOR FILING DATE: 1999-08-06
PRIOR APPLICATION NUMBER: US00/21340
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 67
TYPE: PRT
ORGANISM: Actinobacillus actinomycetemcomitans
US-10-092-243A-15

Query Match 4.0%; Score 11; DB 13; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 TIKGVNAGPE 48
DB 19 TIKGVNAGPE 29

RESULT 13

US-10-282-122A-55826
; Sequence 55826, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55826
LENGTH: 271
TYPE: PRF
ORGANISM: Enterobacter cloacae
US-10-282-122A-55826

Query Match 4.0%; Score 11; DB 12; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKDSRYN 234
DB 219 FVEDKDSRYN 229

RESULT 14

US-10-282-122A-76195
; Sequence 76195, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 76195
LENGTH: 271
TYPE: PRF
ORGANISM: Salmonella typhi
US-10-282-122A-76195

Query Match 4.0%; Score 11; DB 12; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKDSRYN 234
DB 219 FVEDKDSRYN 229

RESULT 15

US-10-282-122A-44792
; Sequence 44792, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

```
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 44792
/ LENGTH: 276
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
US-10-282-122A-44792

Query Match
Best Local Similarity 3.6%; Score 10; DB 12; Length 276;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 ARAIDVDVLA 204
DB 190 ARAIDVDVLA 199

RESULT 16
/ Sequence 44788, Application US/10282122A
/ Publication No. US20040029128A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 44792
/ LENGTH: 276
/ TYPE: PRT
```

```
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 44788
/ LENGTH: 241
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
US-10-282-122A-44788

Query Match
Best Local Similarity 3.3%; Score 9; DB 12; Length 241;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 NARAILLL 156
DB 109 NARAILLL 117

RESULT 17
/ Sequence 61127, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 61127
/ LENGTH: 256
/ TYPE: PRT
```

```
/ ORGANISM: Legionella pneumophila
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (227)..(227)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (241)..(241)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (244)..(244)
/ OTHER INFORMATION: X=any amino acid
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (248)..(248)
/ OTHER INFORMATION: X=any amino acid
US-10-282-122A-61127

Query Match      3.3%  Score 9;  DB 12;  Length 256;
Best Local Similarity 100.0%;  Pred. No. 3.4;
Matches 9;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      140  IAVNDPSN 148
      |||||
Db      125  IAVNDPSN 133

RESULT 18
US-10-282-122A-51371
/ Sequence 51371, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 51371
```

```
/ LENGTH: 262
/ TYPE: PRT
/ ORGANISM: Bordetella pertussis
US-10-282-122A-51371

Query Match      3.3%  Score 9;  DB 12;  Length 262;
Best Local Similarity 100.0%;  Pred. No. 3.5;
Matches 9;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      133  ELKDGATTA 141
      |||||
Db      119  ELKDGATTA 127

RESULT 19
US-10-282-122A-55804
/ Sequence 55804, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 55804
/ LENGTH: 270
/ TYPE: PRT
/ ORGANISM: Enterobacter cloacae
US-10-282-122A-55804

Query Match      3.3%  Score 9;  DB 12;  Length 270;
Best Local Similarity 100.0%;  Pred. No. 3.6;
Matches 9;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      176  DIVENPKXL 184
      |||||
Db      165  DIVENPKXL 173
```

```

RESULT 20
US-10-282-122A-60558
; Sequence 60558, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60558
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60558

Query Match      3.3%; Score 9, DB 12, Length 273;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      140 IAVNDPSN 148
Db      136 IAVNDPSN 144

RESULT 21
US-10-282-122A-77174
; Sequence 77174, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

```

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77174
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77174

Query Match      3.3%; Score 9, DB 12, Length 275;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      112 IVGNTFYVP 120
Db      111 IVGNTFYVP 119

RESULT 22
US-09-764-853-805
; Sequence 805, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT FILING DATE: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Remaining Prior Application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 805
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (20)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

```

LOCATION: (74)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (90)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-805

Query Match 2.9%; Score 8; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 146 PSNLARAL 153
Db 76 PSNLARAL 83

RESULT 23
US-10-091-438-162
Sequence 162, Application US/10091438
Publication No. US2003007760661
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P1217C1
CURRENT APPLICATION NUMBER: US/10/091,438
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 09/764, 879
PRIOR FILING DATE: 2001-01-17, 065
PRIOR APPLICATION NUMBER: 60/179, 065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180, 628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214, 886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217, 487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225, 758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220, 963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217, 496
PRIOR FILING DATE: 2000-07-11, 447
PRIOR APPLICATION NUMBER: 60/225, 447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218, 290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225, 757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226, 868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216, 647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225, 267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216, 880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225, 270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251, 869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235, 834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234, 274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234, 223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228, 924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224, 518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236, 369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224, 519

PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220, 964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241, 809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249, 299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236, 327
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/241, 785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244, 617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225, 268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236, 368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251, 856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251, 868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229, 344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234, 997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229, 343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231, 413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229, 509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236, 367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237, 039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237, 038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237, 040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240, 960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239, 935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239, 937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241, 787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246, 474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246, 532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249, 216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226, 681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225, 759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225, 213
PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08

Query Match 2.9%; Score 8; DB 14; Length 96;
Best Local Similarity 100.0%; Pred. No. 13;
Matches: 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 146 PSNLARL 153
DB 76 PSNLARL 83

RESULT 24
US-10-424-599-262926
Sequence 262926, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 262926
LENGTH: 143
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_79443C.1.pgp
US-10-424-599-262926

Query Match 2.9%; Score 8; DB 12; Length 143;
Best Local Similarity 100.0%; Pred. No. 18;
Matches: 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 188 EVDTSVAA 195
DB 93 EVDTSVAA 100

RESULT 25
US-09-815-242-11656
Sequence 11656, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl U.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16

```
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11656
/ LENGTH: 240
/ TYPE: PRT
/ ORGANISM: Klebsiella pneumoniae
US-09-815-242-11656

Query Match      2.9%; Score 8; DB 9; Length 240;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 VGNFVVP 120
DB      75 VGNFVVP 82

RESULT 26
US-10-282-122A-52976
/ Sequence 52976, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zysek, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 52976
/ LENGTH: 263
/ TYPE: PRT
/ ORGANISM: Clostridium difficile
US-10-282-122A-52976

Query Match      2.9%; Score 8; DB 12; Length 263;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      229 DSPYVNI 236
DB      111 DSPYVNI 111
```

```
DB      216 DSPYVNI 223

RESULT 27
US-10-282-122A-51201
/ Sequence 51201, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zysek, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 51201
/ LENGTH: 265
/ TYPE: PRT
/ ORGANISM: Bordetella pertussis
US-10-282-122A-51201

Query Match      2.9%; Score 8; DB 12; Length 265;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      131 LNEIKDGA 138
DB      120 LNEIKDGA 127

RESULT 28
US-10-282-122A-49339
/ Sequence 49339, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zysek, Judith
```

```

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49339
; LENGTH: 270
; TYPE: PRF
; ORGANISM: Burkholderia fungorum
; US-10-282-122A-49339

Query Match      2.9%; Score 8; DB 12; Length 270;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Cy      141 AVNDPSN 148
Db      132 AVNDPSN 139

RESULT 29
; US-09-815-242-10050
; Sequence 10050, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyckind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43291
; LENGTH: 271
; TYPE: PRF
; ORGANISM: Escherichia coli
; US-10-282-122A-43291
```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10050
; LENGTH: 271
; TYPE: PRF
; ORGANISM: Escherichia coli
; US-09-815-242-10050

Query Match      2.9%; Score 8; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Cy      113 VGNFTVYP 120
Db      106 VGNFTVYP 113

RESULT 30
; US-10-282-122A-43291
; Sequence 43291, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43291
; LENGTH: 271
; TYPE: PRF
; ORGANISM: Escherichia coli
; US-10-282-122A-43291
```

Query Match 2.9%; Score 8; DB 12; Length 271;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 VGNTEVP 120
Db 106 VGNTEVP 113

RESULT 31
US-10-282-122A-59600
; Sequence 59600, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59600
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59600

Query Match 2.9%; Score 8; DB 12; Length 271;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 VGNTEVP 120
Db 106 VGNTEVP 113

RESULT 32
US-10-282-122A-74389
; Sequence 74389, Application US/10282122A
; Publication No. US20040029129A1

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74389
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74389

Query Match 2.9%; Score 8; DB 12; Length 281;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYQDEV 260
Db 258 KAYQDEV 265

RESULT 33
US-10-282-122A-57911
; Sequence 57911, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 57911
/ LENGTH: 282
/ TYPE: PRT
/ ORGANISM: Enterococcus faecium
/ US-10-282-122A-57911
```

```
Query Match      2.9%; Score 8; DB 12; Length 282;
Best Local Similarity 100.0%; Pred.No.35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      133 ELKDGTI 140
Db      131 ELKDGTI 138
```

```
RESULT 34
US-10-282-122A-47704
/ Sequence 47704, Application US/10282122A
/ Publication No. US2004002919A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
```

```
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 47704
/ LENGTH: 295
/ TYPE: PRT
/ ORGANISM: Burkholderia cepacia
/ US-10-282-122A-47704
```

```
Query Match      2.9%; Score 8; DB 12; Length 295;
Best Local Similarity 100.0%; Pred.No.36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      141 AVNDPSN 148
Db      157 AVNDPSN 164
```

```
RESULT 35
US-09-764-853-621
/ Sequence 621, Application US/09764853
/ Patent No. US20020090672A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: P206
/ CURRENT APPLICATION NUMBER: US/09/764,853
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - consult PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 939
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 621
/ LENGTH: 379
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-764-853-621
```

```
Query Match      2.9%; Score 8; DB 9; Length 379;
Best Local Similarity 100.0%; Pred.No.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      146 PSNLARAL 153
Db      128 PSNLARAL 135
```

```
RESULT 36
US-10-369-493-20865
/ Sequence 20865, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
```

SEQ ID NO 20865
LENGTH: 467
TYPE: PRT
ORGANISM: Rhodospirillum rubrum

Query Match
Best Local Similarity 100.0%; Score 8; DB 15; Length 467;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 KDGATIAV 142
DB 336 KDGATIAV 343

RESULT 37
US-10-369-493-15592
Sequence 15592, Application US/10369493
Patent No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT FILING DATE: 2003-02-28
PRIORITY FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15592
LENGTH: 641
TYPE: PRT
ORGANISM: Xanthomonas campestris
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(641)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-15592

Query Match
Best Local Similarity 100.0%; Score 8; DB 15; Length 641;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IALGCSN 23
DB 2 IALGCSN 9

RESULT 38
US-09-854-845-45
Sequence 45, Application US/09854845
Patent No. US20020098491A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotide
FILE REFERENCE: LEX-0177-USA
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-18
PRIORITY FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 45
LENGTH: 739

TYPE: PRT
ORGANISM: homo sapiens
US-09-854-845-45

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 739;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 PSNLARAL 153
DB 488 PSNLARAL 495

RESULT 39
US-09-854-845-43
Sequence 43, Application US/09854845
Patent No. US20020098491A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotide
FILE REFERENCE: LEX-0177-USA
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-18
PRIORITY FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 744
TYPE: PRT
ORGANISM: homo sapiens
US-09-854-845-43

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 744;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 PSNLARAL 153
DB 493 PSNLARAL 500

RESULT 40
US-09-854-845-49
Sequence 49, Application US/09854845
Patent No. US20020098491A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotide
FILE REFERENCE: LEX-0177-USA
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-18
PRIORITY FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49
LENGTH: 766
TYPE: PRT
ORGANISM: homo sapiens
US-09-854-845-49

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 766;

Tue Jun 22 11:04:02 2004

us-10-018-672-2.oligo.rapb

Page 23

Best Local Similarity 100.0%; Pred. No. 89;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 146 PSNLRAL 153
Db 488 PSNLRAL 495

Search completed: June 16, 2004, 11:23:05
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:12:42 ; Search time 45 Seconds
(without alignments)
1935.178 Million cell updates/sec

Title: US-10-018-672-2

Perfect score: 276
Sequence: 1 MNFGKINGICALASGIALAG.....TDEVEAKKQKQFDGVIKGM 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_rhiz: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeoph: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	16	5.8	270	Q7VM93 haemophilus
2	11	4.0	271	Q83MC6 shigella fl
3	10	3.6	271	Q7WIF5 bordetella
4	10	3.6	271	Q7WIF6 bordetella
5	10	3.6	281	Q7W613 bordetella
6	10	3.3	259	Q69441 bordetella
7	9	3.3	262	Q7W613 bordetella
8	9	3.3	262	Q7W613 bordetella
9	9	3.3	262	Q7W613 bordetella
10	9	3.3	262	Q7W613 bordetella
11	9	3.3	273	Q8YA74 listeria mo
12	9	3.3	276	Q7WML8 bordetella
13	9	3.3	276	Q7WML8 bordetella
14	9	3.3	276	Q7WML8 bordetella
15	9	3.3	278	Q8YD39 bruceella me
16	8	2.9	89	Q8YV86 bruceella su
				Q9UYT7 pyrococcus

17	8	2.9	255	Q9AG13 pseudomonas
18	8	2.9	259	Q7XSM6 pseudomonas
19	8	2.9	259	Q8XNL7 clostridium
20	8	2.9	259	Q8U7G0 bordetella
21	8	2.9	265	Q7W613 bordetella
22	8	2.9	265	Q7W613 bordetella
23	8	2.9	265	Q7W613 bordetella
24	8	2.9	266	Q8P4S9 xanthomonas
25	8	2.9	269	Q7V195 helicobacte
26	8	2.9	269	Q8P6F0 xanthomonas
27	8	2.9	269	Q8P6F0 xanthomonas
28	8	2.9	270	Q8W613 bordetella
29	8	2.9	272	Q8W613 bordetella
30	8	2.9	272	Q8W613 bordetella
31	8	2.9	277	Q7Z3D8 bordetella
32	8	2.9	281	Q8P4E4 bordetella
33	8	2.9	281	Q8P4E4 bordetella
34	8	2.9	305	Q8P4E4 bordetella
35	8	2.9	305	Q8P4E4 bordetella
36	8	2.9	351	Q8ZSM1 bordetella
37	8	2.9	351	Q8ZSM1 bordetella
38	8	2.9	374	Q8ZSM1 bordetella
39	8	2.9	446	Q8P4E4 bordetella
40	8	2.9	512	Q8P4E4 bordetella
41	8	2.9	556	Q8P4E4 bordetella
42	8	2.9	652	Q8P4E4 bordetella
43	8	2.9	652	Q8P4E4 bordetella
44	8	2.9	761	Q8P4E4 bordetella
45	8	2.9	1220	Q8P4E4 bordetella
46	8	2.9	1406	Q8P4E4 bordetella
47	8	2.9	3665	Q8P4E4 bordetella
48	8	2.5	27	Q8P4E4 bordetella
49	8	2.5	47	Q8P4E4 bordetella
50	8	2.5	68	Q8P4E4 bordetella
51	8	2.5	74	Q8P4E4 bordetella
52	8	2.5	87	Q8P4E4 bordetella
53	8	2.5	92	Q8P4E4 bordetella
54	8	2.5	104	Q8P4E4 bordetella
55	8	2.5	105	Q8P4E4 bordetella
56	8	2.5	109	Q8P4E4 bordetella
57	8	2.5	114	Q8P4E4 bordetella
58	8	2.5	114	Q8P4E4 bordetella
59	8	2.5	114	Q8P4E4 bordetella
60	8	2.5	114	Q8P4E4 bordetella
61	8	2.5	115	Q8P4E4 bordetella
62	8	2.5	115	Q8P4E4 bordetella
63	8	2.5	121	Q8P4E4 bordetella
64	8	2.5	123	Q8P4E4 bordetella
65	8	2.5	123	Q8P4E4 bordetella
66	8	2.5	134	Q8P4E4 bordetella
67	8	2.5	137	Q8P4E4 bordetella
68	8	2.5	138	Q8P4E4 bordetella
69	8	2.5	142	Q8P4E4 bordetella
70	8	2.5	150	Q8P4E4 bordetella
71	8	2.5	151	Q8P4E4 bordetella
72	8	2.5	152	Q8P4E4 bordetella
73	8	2.5	154	Q8P4E4 bordetella
74	8	2.5	156	Q8P4E4 bordetella
75	8	2.5	161	Q8P4E4 bordetella
76	8	2.5	161	Q8P4E4 bordetella
77	8	2.5	163	Q8P4E4 bordetella
78	8	2.5	163	Q8P4E4 bordetella
79	8	2.5	170	Q8P4E4 bordetella
80	8	2.5	174	Q8P4E4 bordetella
81	8	2.5	176	Q8P4E4 bordetella
82	8	2.5	178	Q8P4E4 bordetella
83	8	2.5	185	Q8P4E4 bordetella
84	8	2.5	185	Q8P4E4 bordetella
85	8	2.5	188	Q8P4E4 bordetella
86	8	2.5	188	Q8P4E4 bordetella
87	8	2.5	196	Q8P4E4 bordetella
88	8	2.5	196	Q8P4E4 bordetella
89	8	2.5	199	Q8P4E4 bordetella

90	2.5	207	3	Q9URD7	Q9URD7 aspergillus	13	7	2.5	300	16	Q8UDP3	Q8UDP3 agrobacteri
91	2.5	207	16	Q989P3	Q989P3 rhizobium 1	14	7	2.5	300	15	Q989V1	Q989V1 deinococcus
92	2.5	209	16	Q8KGA0	Q8KGA0 chlorobium	15	7	2.5	300	17	Q971G8	Q971G8 sulfolobum
93	2.5	209	16	Q89UK1	Q89UK1 bradyrhizob	16	7	2.5	303	2	Q932A6	Q932A6 rhizobium 1
94	2.5	215	10	Q49911	Q49911 nicotiana t	167	7	2.5	303	16	Q8CVA8	Q8CVA8 oceanobacil
95	2.5	218	16	Q89120	Q89120 bradyrhizob	168	7	2.5	306	16	Q8U782	Q8U782 agrobacteri
96	2.5	220	16	Q889Q8	Q889Q8 pseudomonas	169	7	2.5	307	8	Q9XM22	Q9XM22 ascaris suu
97	2.5	224	17	Q973Z5	Q973Z5 sulfolobus	170	7	2.5	307	16	Q9CHT8	Q9CHT8 lactococcus
98	2.5	231	16	Q83HV9	Q83HV9 tripheryma	171	7	2.5	317	13	Q91336	Q91336 rana sylvat
99	2.5	231	16	Q83GCO	Q83GCO nicotiana t	172	7	2.5	322	16	Q89HG6	Q89HG6 bradyrhizob
100	2.5	232	10	Q49910	Q49910 nicotiana t	173	7	2.5	323	17	Q8EXY1	Q8EXY1 methanosarc
101	2.5	232	10	Q9MWU3	Q9MWU3 nicotiana t	174	7	2.5	335	17	Q8U752	Q8U752 methanosarc
102	2.5	232	10	Q84LC7	Q84LC7 chrysanthem	175	7	2.5	341	5	Q9NB46	Q9NB46 caenorhabdi
103	2.5	234	16	Q9ZLC9	Q9ZLC9 heliobacter	176	7	2.5	344	16	Q67697	Q67697 aquifex aeo
104	2.5	239	16	Q9PMD0	Q9PMD0 campylobact	177	7	2.5	350	16	Q87CL9	Q87CL9 xylella fas
105	2.5	241	16	Q9A5P2	Q9A5P2 callobacter	178	7	2.5	352	16	Q8Y4K6	Q8Y4K6 escherichia
106	2.5	242	16	Q83CE8	Q83CE8 bradyrhizob	179	7	2.5	355	4	Q8NCU3	Q8NCU3 homo sapien
107	2.5	244	16	Q98D84	Q98D84 rhizobium 1	180	7	2.5	358	16	Q7WN65	Q7WN65 bordetella
108	2.5	247	16	Q98P84	Q98P84 campylobact	181	7	2.5	363	16	Q8THV7	Q8THV7 methanosarc
109	2.5	254	16	Q864Y2	Q864Y2 bifidobacte	182	7	2.5	363	16	Q8PPU3	Q8PPU3 xanthomonas
110	2.5	256	16	Q88XL7	Q88XL7 pseudomonas	183	7	2.5	380	13	Q42317	Q42317 cyprinus ca
111	2.5	260	16	Q9H768	Q9H768 vibrio para	184	7	2.5	380	16	Q9PCG2	Q9PCG2 xylella fas
112	2.5	260	16	Q870I2	Q870I2 vibrio para	185	7	2.5	386	17	P95972	P95972 campylobact
113	2.5	261	16	Q9PF05	Q9PF05 xylella fas	186	7	2.5	397	16	Q9AA49	Q9AA49 vibrio chol
114	2.5	261	16	Q8D8T3	Q8D8T3 vibrio vuln	187	7	2.5	400	16	Q9KN80	Q9KN80 vibrio chol
115	2.5	261	16	Q87ML7	Q87ML7 xylella fas	188	7	2.5	401	2	Q51696	Q51696 bacteroides
116	2.5	262	16	Q9KRF8	Q9KRF8 vibrio chol	189	7	2.5	402	16	Q8A2Z7	Q8A2Z7 streptomyces
117	2.5	263	16	Q9KMS9	Q9KMS9 vibrio algi	190	7	2.5	404	16	Q9WKZ7	Q9WKZ7 salmonella
118	2.5	264	16	Q7VMV2	Q7VMV2 haemophilus	191	7	2.5	406	16	Q8ZMW4	Q8ZMW4 salmonella
119	2.5	265	16	Q83VH9	Q83VH9 treponema h	192	7	2.5	406	16	Q8ZAI6	Q8ZAI6 salmonella
120	2.5	265	16	Q9CMB7	Q9CMB7 pasteurella	193	7	2.5	410	2	Q8G1S7	Q8G1S7 synechococc
121	2.5	266	16	Q827C4	Q827C4 salmonella	194	7	2.5	411	10	Q8RVC7	Q8RVC7 chlamydomon
122	2.5	268	16	Q8YAR0	Q8YAR0 escherichia	195	7	2.5	420	16	Q8XY93	Q8XY93 ralestonia s
123	2.5	268	16	Q83T26	Q83T26 salmonella	196	7	2.5	420	16	Q06832	Q06832 mycobacteri
124	2.5	269	16	Q86191	Q86191 erwania chr	197	7	2.5	424	16	Q7U038	Q7U038 mycobacteri
125	2.5	269	16	Q8DFC1	Q8DFC1 vibrio vuln	198	7	2.5	424	16	Q8NSE9	Q8NSE9 corynebacte
126	2.5	270	16	Q8G8H9	Q8G8H9 erwania pyr	199	7	2.5	428	16	Q8XY06	Q8XY06 ralestonia s
127	2.5	270	16	Q81VM1	Q81VM1 bacillus an	200	7	2.5	432	16	Q8U1U0	Q8U1U0 agrobacteri
128	2.5	270	16	Q81I25	Q81I25 bacillus ce	201	7	2.5	432	16	Q9RT67	Q9RT67 denirococcus
129	2.5	271	16	Q26084	Q26084 helicobacte	202	7	2.5	434	16	Q9RT67	Q9RT67 chlamydomon
130	2.5	271	16	Q92X45	Q92X45 helicobacte	203	7	2.5	454	16	Q8RA18	Q8RA18 chlamydomon
131	2.5	271	16	Q82E27	Q82E27 yerisina pe	204	7	2.5	455	10	Q94IP6	Q94IP6 anguilla an
132	2.5	273	16	Q92E25	Q92E25 listeria in	205	7	2.5	458	13	Q98843	Q98843 anguilla an
133	2.5	274	2	Q8KUB0	Q8KUB0 vibrio fisc	206	7	2.5	460	8	Q9AS89	Q9AS89 corynebacte
134	2.5	274	2	Q7VMX6	Q7VMX6 borrelia ga	207	7	2.5	460	16	Q8NSI8	Q8NSI8 corynebacte
135	2.5	275	16	Q8ZG10	Q8ZG10 yerisina pe	208	7	2.5	468	5	Q19775	Q19775 caenorhabdi
136	2.5	276	16	Q82R89	Q82R89 streptomyce	209	7	2.5	477	4	Q8NEQ8	Q8NEQ8 homo sapien
137	2.5	276	16	Q928M0	Q928M0 listeria in	210	7	2.5	478	11	Q8BLD9	Q8BLD9 mus musculu
138	2.5	276	16	Q8Y4M0	Q8Y4M0 listeria mo	211	7	2.5	484	16	Q98HCO	Q98HCO rhizobium 1
139	2.5	279	16	Q8ENU4	Q8ENU4 oceanobacil	212	7	2.5	496	16	Q8CMY1	Q8CMY1 staphylococ
140	2.5	282	16	Q8HMD2	Q8HMD2 plasmodium	213	7	2.5	523	16	Q8XOM7	Q8XOM7 ralestonia s
141	2.5	282	16	Q89HG5	Q89HG5 bradyrhizob	214	7	2.5	527	5	Q9VC35	Q9VC35 drosophila
142	2.5	284	2	Q93EA5	Q93EA5 rhizobium 1	215	7	2.5	527	5	Q86PC4	Q86PC4 drosophila
143	2.5	284	16	Q98DA1	Q98DA1 rhizobium 1	216	7	2.5	527	5	Q9XZ04	Q9XZ04 drosophila
144	2.5	286	16	Q9CIN8	Q9CIN8 lactococcus	217	7	2.5	529	17	Q8PYV7	Q8PYV7 methanosarc
145	2.5	286	16	Q9CIN7	Q9CIN7 lactococcus	218	7	2.5	530	16	Q9KP41	Q9KP41 vibrio chol
146	2.5	287	2	Q9JPI7	Q9JPI7 neisseria m	219	7	2.5	541	10	Q84TG1	Q84TG1 arabidopsis
147	2.5	287	2	Q9JPI7	Q9JPI7 neisseria m	220	7	2.5	542	5	Q8WZ86	Q8WZ86 drosophila
148	2.5	289	16	Q82N55	Q82N55 streptomyce	221	7	2.5	545	10	Q9G8J1	Q9G8J1 arabidopsis
149	2.5	290	13	Q800B7	Q800B7 rana rugosa	222	7	2.5	549	16	Q97M37	Q97M37 clostridium
150	2.5	290	13	Q800B6	Q800B6 rana rugosa	223	7	2.5	554	5	Q91774	Q91774 caenorhabdi
151	2.5	290	13	Q7ZT24	Q7ZT24 rana rugosa	224	7	2.5	557	10	Q8VZB2	Q8VZB2 arabidopsis
152	2.5	291	3	Q8C445	Q8C445 penicillium	225	7	2.5	563	16	Q891U9	Q891U9 clostridium
153	2.5	291	3	Q7YSW6	Q7YSW6 bacterioph	226	7	2.5	566	5	Q810H9	Q810H9 drosophila
154	2.5	296	16	Q7UOG5	Q7UOG5 rhodospirill	227	7	2.5	570	16	Q9S1S8	Q9S1S8 streptomyce
155	2.5	298	13	Q8SOH5	Q8SOH5 bos taurus	228	7	2.5	580	16	Q9S856	Q9S856 synechocyst
156	2.5	298	13	Q9YTC4	Q9YTC4 rana rugosa	229	7	2.5	581	16	Q89NWI	Q89NWI bradyrhizob
157	2.5	298	13	Q92RH2	Q92RH2 rana rugosa	230	7	2.5	585	16	Q8A7X1	Q8A7X1 bacteroides
158	2.5	298	13	Q9PRH1	Q9PRH1 rana rugosa	231	7	2.5	586	16	Q97UT5	Q97UT5 clostridium
159	2.5	298	13	Q8AYW3	Q8AYW3 xenopus lae	232	7	2.5	598	5	Q8PII0	Q8PII0 drosophila
160	2.5	298	13	Q919W9	Q919W9 xenopus lae	233	7	2.5	599	5	Q25310	Q25310 leishmania
161	2.5	298	13	Q8JHI0	Q8JHI0 brachydanio	234	7	2.5	600	5	Q01530	Q01530 trypanosoma
162	2.5	300	5	Q01813	Q01813 caenorhabdi	235	7	2.5	601	16	Q8AB55	Q8AB55 bacteroides

236	7	2.5	610	5	Q19848	Q19848 caenorhabdi	309	7	2.5	18412	13	Q7Z261	Q7Z261 brachydanio
237	7	2.5	610	10	Q9L7N3	Q9L7N3 arabidopsis	310	6	2.2	20	2	Q9R4K6	Q9R4K6 staphylococ
238	7	2.5	611	17	Q9UZU0	Q9UZU0 pyrococcus	311	6	2.2	20	15	Q9DYR0	Q9DYR0 human immun
239	7	2.5	616	16	Q9XQK4	Q9XQK4 vibrio chol	312	6	2.2	33	4	Q16148	Q16148 homo sapien
240	7	2.5	634	5	Q8STZ1	Q8STZ1 encephalito	313	6	2.2	35	16	Q8CQ13	Q8CQ13 staphylococ
241	7	2.5	634	5	Q8ST44	Q8ST44 encephalito	314	6	2.2	42	10	P93072	P93072 bambusa sp.
242	7	2.5	638	16	Q8RT72	Q8RT72 fusobacteri	315	6	2.2	44	4	Q13013	Q13013 homo sapien
243	7	2.5	652	16	Q97OT2	Q97OT2 streptococ	316	6	2.2	44	16	Q8CSY2	Q8CSY2 staphylococ
244	7	2.5	652	16	Q8DP89	Q8DP89 streptococ	317	6	2.2	44	16	Q7WHS7	Q7WHS7 bordetella
245	7	2.5	653	10	Q9LSA0	Q9LSA0 arabidopsis	318	6	2.2	44	16	Q7W6U1	Q7W6U1 bordetella
246	7	2.5	660	16	Q8Y140	Q8Y140 ralteonias	319	6	2.2	45	2	P83309	P83309 comonomas a
247	7	2.5	668	5	Q3V8D2	Q3V8D2 drosophila	320	6	2.2	45	4	Q8TD77	Q8TD77 homo sapien
248	7	2.5	668	5	Q8SZH5	Q8SZH5 drosophila	321	6	2.2	45	4	Q12847	Q12847 homo sapien
249	7	2.5	670	5	Q3VEY1	Q3VEY1 drosophila	322	6	2.2	48	2	Q9X9H7	Q9X9H7 yersinia en
250	7	2.5	670	16	Q51338	Q51338 borrelia bu	323	6	2.2	50	16	Q8X3B6	Q8X3B6 escherichia
251	7	2.5	691	3	Q3URQ2	Q3URQ2 saccharomyc	324	6	2.2	50	16	Q8Z8R0	Q8Z8R0 streptomyce
252	7	2.5	691	16	Q34523	Q34523 helicobacte	325	6	2.2	56	10	Q9LJ00	Q9LJ00 oryza sativ
253	7	2.5	692	16	Q8CY25	Q8CY25 corynebacte	326	6	2.2	57	9	Q8SDJ4	Q8SDJ4 streptococ
254	7	2.5	694	16	Q9X120	Q9X120 thermocoga	327	6	2.2	57	10	Q7X1N8	Q7X1N8 oryza sativ
255	7	2.5	696	16	Q92VD4	Q92VD4 rhizobium m	328	6	2.2	60	2	Q9X127	Q9X127 edwardsiella
256	7	2.5	703	16	Q7UXT8	Q7UXT8 rhodospirell	329	6	2.2	60	3	Q17711	Q17711 trichoderma
257	7	2.5	718	4	Q9HCD1	Q9HCD1 homo sapien	330	6	2.2	62	16	Q89VU7	Q89VU7 bradyrhizob
258	7	2.5	721	5	Q9VMU5	Q9VMU5 drosophila	331	6	2.2	62	16	Q928D7	Q928D7 yersinia in
259	7	2.5	721	10	Q941P7	Q941P7 chlamydomon	332	6	2.2	64	16	Q9S1E1	Q9S1E1 yersinia pe
260	7	2.5	723	16	Q8DA43	Q8DA43 vibrio vuln	333	6	2.2	66	2	Q9AP90	Q9AP90 uncultured
261	7	2.5	734	12	Q8OY51	Q8OY51 heterobasid	334	6	2.2	67	2	Q85232	Q85232 staphylococ
262	7	2.5	741	16	Q8E1V5	Q8E1V5 shewanella	335	6	2.2	67	5	Q20433	Q20433 caenorhabdi
263	7	2.5	749	5	Q9W4E0	Q9W4E0 drosophila	336	6	2.2	67	10	Q852U3	Q852U3 brassica ju
264	7	2.5	762	17	Q976U7	Q976U7 sulfolobus	337	6	2.2	68	16	Q83E18	Q83E18 coxiella bu
265	7	2.5	776	2	Q8RN57	Q8RN57 campylobact	338	6	2.2	67	16	Q9RRY6	Q9RRY6 deinococcus
266	7	2.5	796	4	Q8NC87	Q8NC87 homo sapien	339	6	2.2	68	16	Q9KEB4	Q9KEB4 bacillus ha
267	7	2.5	796	5	Q86NX5	Q86NX5 drosophila	340	6	2.2	68	16	Q8FRCS	Q8FRCS corynebacte
268	7	2.5	800	2	Q8RN16	Q8RN16 campylobact	341	6	2.2	69	2	Q52023	Q52023 staphylococ
269	7	2.5	812	16	Q8FWG2	Q8FWG2 corynebacte	342	6	2.2	70	10	Q7XJ07	Q7XJ07 oryza sativ
270	7	2.5	860	16	Q8NM83	Q8NM83 corynebacte	343	6	2.2	70	16	Q8ZDW4	Q8ZDW4 yersinia pe
271	7	2.5	867	2	Q7WZ88	Q7WZ88 nonomurea	344	6	2.2	71	7	Q9GJ66	Q9GJ66 salmo trutt
272	7	2.5	875	16	Q8Y6W6	Q8Y6W6 listeria mo	345	6	2.2	71	7	Q9GJ63	Q9GJ63 salmo trutt
273	7	2.5	905	16	Q8UCC1	Q8UCC1 agrobacteri	346	6	2.2	71	7	Q9GJ65	Q9GJ65 salmo trutt
274	7	2.5	913	16	Q894I2	Q894I2 clostridium	347	6	2.2	71	7	Q9GJ64	Q9GJ64 salmo trutt
275	7	2.5	933	16	Q83GL1	Q83GL1 tropheryma	348	6	2.2	71	16	Q87NS5	Q87NS5 vibrio para
276	7	2.5	938	10	Q9AM83	Q9AM83 gulliardia	349	6	2.2	72	12	Q52023	Q52023 staphylococ
277	7	2.5	939	16	Q83HL6	Q83HL6 tropheryma	350	6	2.2	72	12	Q8Q0Q9	Q8Q0Q9 camelipox vi
278	7	2.5	965	5	Q8WMZ6	Q8WMZ6 cryptospori	351	6	2.2	73	16	Q68801	Q68801 pseudomonas
279	7	2.5	967	16	Q98GNS	Q98GNS rhizobium l	352	6	2.2	73	16	Q881L6	Q881L6 pseudomonas
280	7	2.5	1012	10	Q7XLU5	Q7XLU5 oryza sativ	353	6	2.2	74	17	Q8ZUW3	Q8ZUW3 pyrobaculum
281	7	2.5	1036	16	Q9KYL2	Q9KYL2 vibrio chol	354	6	2.2	76	9	Q8H9X8	Q8H9X8 pseudomonas
282	7	2.5	1039	16	Q8DDC6	Q8DDC6 vibrio vuln	355	6	2.2	76	10	Q9LFR0	Q9LFR0 arabidopsis
283	7	2.5	1040	16	Q87NM1	Q87NM1 arabidopsis	356	6	2.2	79	12	Q8Q078	Q8Q078 camelipox vi
284	7	2.5	1095	10	Q9CTH5	Q9CTH5 arabidopsis	357	6	2.2	81	2	Q8S4L1	Q8S4L1 streptococ
285	7	2.5	1118	16	Q97E71	Q97E71 clostridium	358	6	2.2	82	17	Q8TRV0	Q8TRV0 methanopyru
286	7	2.5	1130	5	Q9X119	Q9X119 caenorhabdi	359	6	2.2	82	16	Q8H4U5	Q8H4U5 bos taurus
287	7	2.5	1130	16	Q7UT96	Q7UT96 rhodospirell	360	6	2.2	83	6	Q864U5	Q864U5 escherichia
288	7	2.5	1171	16	Q8D6T4	Q8D6T4 vibrio vuln	361	6	2.2	83	16	Q8XP54	Q8XP54 clostridium
289	7	2.5	1233	16	Q8NSE4	Q8NSE4 corynebacte	362	6	2.2	85	12	Q8V6P2	Q8V6P2 halovirus h
290	7	2.5	1240	4	Q8NFM2	Q8NFM2 homo sapien	363	6	2.2	85	16	Q7TD11	Q7TD11 halovirus h
291	7	2.5	1283	10	Q8L1W4	Q8L1W4 oryza sativ	364	6	2.2	85	12	Q9JZX3	Q9JZX3 neisseria m
292	7	2.5	1333	5	Q24262	Q24262 drosophila	365	6	2.2	86	3	Q9US10	Q9US10 escherichia
293	7	2.5	1431	10	Q7XV44	Q7XV44 oryza sativ	366	6	2.2	86	16	Q8DIM5	Q8DIM5 synecococ
294	7	2.5	1496	2	Q93IE8	Q93IE8 actinobacil	367	6	2.2	87	10	Q8L9S3	Q8L9S3 arabidopsis
295	7	2.5	1498	2	Q91448	Q91448 arthropacte	368	6	2.2	88	17	Q8ZV74	Q8ZV74 pyrobaculum
296	7	2.5	1615	4	Q9NYS8	Q9NYS8 homo sapien	369	6	2.2	90	16	Q8X5V2	Q8X5V2 escherichia
297	7	2.5	1616	4	Q8NEV4	Q8NEV4 homo sapien	370	6	2.2	90	16	Q8X5V2	Q8X5V2 escherichia
298	7	2.5	1784	5	Q9VES4	Q9VES4 drosophila	371	6	2.2	90	16	Q8FB46	Q8FB46 mycobacteri
299	7	2.5	1795	16	Q9ZUF7	Q9ZUF7 ticketella	372	6	2.2	90	16	Q8XGB5	Q8XGB5 salmonella
300	7	2.5	1919	10	Q8LRK9	Q8LRK9 arabidopsis	373	6	2.2	90	16	Q7TYW0	Q7TYW0 mycobacteri
301	7	2.5	1994	10	Q9LPI9	Q9LPI9 arabidopsis	374	6	2.2	91	16	Q83B37	Q83B37 coxiella bu
302	7	2.5	2147	2	Q91350	Q91350 pseudomonas	375	6	2.2	92	2	Q66391	Q66391 methylomona
303	7	2.5	2154	16	Q9HVG6	Q9HVG6 pseudomonas	376	6	2.2	92	5	Q18742	Q18742 caenorhabdi
304	7	2.5	4066	2	Q8XLL5	Q8XLL5 streptomyce	377	6	2.2	92	11	Q8BLI4	Q8BLI4 mus musculu
305	7	2.5	4340	2	Q30764	Q30764 streptomyce	378	6	2.2	93	16	Q7VIN3	Q7VIN3 helicobacte
306	7	2.5	4688	16	Q9PQ08	Q9PQ08 ureaplasma	379	6	2.2	93	17	Q26421	Q26421 methanobact
307	7	2.5	6310	16	Q88FP2	Q88FP2 pseudomonas	380	6	2.2	95	11	Q8K2K1	Q8K2K1 mus musculu
308	7	2.5	6396	2	Q9KID7	Q9KID7 streptomyce	381	6	2.2	95	16	Q83JC6	Q83JC6 shigella fl

382	6	2.2	95	16	082WV6	082WV6 nitrosomona
383	6	2.2	96	9	Q37904	Q37904 enterobacte
384	6	2.2	96	11	063156	063156 rattus norv
385	6	2.2	96	16	07V441	07V441 prochloroco
386	6	2.2	97	6	086214	086214 bos taurus
387	6	2.2	97	16	098050	098050 rhizobium 1
388	6	2.2	98	5	001649	001649 drosophila
389	6	2.2	98	6	095N84	095N84 ovis aries
390	6	2.2	99	9	07Y423	07Y423 bacterioph
391	6	2.2	99	16	09K624	09K624 bacillus ha
392	6	2.2	99	16	08YFC2	08YFC2 bruceella me
393	6	2.2	99	16	082363	082363 salmonella
394	6	2.2	101	2	08RJ00	08RJ00 nitrosococ
395	6	2.2	101	16	08DSC9	08DSC9 streptococ
396	6	2.2	101	16	07J314	07J314 prochloroco
397	6	2.2	102	16	08LEU7	08LEU7 arabidopsis
398	6	2.2	102	16	08W453	08W453 arabidopsis
399	6	2.2	102	16	08DM61	08DM61 streptococ
400	6	2.2	102	16	093Y71	093Y71 streptococ
401	6	2.2	103	2	07WUB7	07WUB7 escherichia
402	6	2.2	103	16	07WOC9	07WOC9 candidatus
403	6	2.2	104	16	092DX8	092DX8 listeria in
404	6	2.2	104	16	08Y959	08Y959 listeria mo
405	6	2.2	105	6	093R83	093R83 vibrio fisc
406	6	2.2	105	6	09TSTY3	09TSTY3 sus scrofa
407	6	2.2	105	16	07WE14	07WE14 bordetella
408	6	2.2	105	16	07W366	07W366 bordetella
409	6	2.2	105	16	07VSO1	07VSO1 bordetella
410	6	2.2	106	10	093580	093580 stipa virid
411	6	2.2	106	10	09M1Y4	09M1Y4 arabidopsis
412	6	2.2	107	16	088VP4	088VP4 lactobacill
413	6	2.2	108	16	08XK16	08XK16 ralstonia s
414	6	2.2	108	16	08EM18	08EM18 oceanobacil
415	6	2.2	108	16	08D9H7	08D9H7 vibrio vuln
416	6	2.2	109	2	09EUV9	09EUV9 listeria in
417	6	2.2	109	16	08YBZ2	08YBZ2 bruceella me
418	6	2.2	110	5	09YUO0	09YUO0 drosophila
419	6	2.2	110	16	09K6V9	09K6V9 bacillus ha
420	6	2.2	110	16	09F5L4	09F5L4 brachyriazob
421	6	2.2	111	10	080578	080578 arabidopsis
422	6	2.2	111	16	08NUL6	08NUL6 corynebacte
423	6	2.2	112	16	096695	096695 bacillus su
424	6	2.2	112	16	08FST0	08FST0 corynebacte
425	6	2.2	112	16	07USG0	07USG0 rhodospirill
426	6	2.2	113	3	060048	060048 pleurotus o
427	6	2.2	113	11	09D4T7	09D4T7 mus musculu
428	6	2.2	113	17	08ZPB5	08ZPB5 pyrobaculum
429	6	2.2	114	9	0856B1	0856B1 mycobacteri
430	6	2.2	114	11	08CAK0	08CAK0 mus musculu
431	6	2.2	114	12	099M00	099M00 mus musculu
432	6	2.2	115	5	09EMM7	09EMM7 amastrea moo
433	6	2.2	115	5	017259	017259 brachionus
434	6	2.2	115	5	09VAK8	09VAK8 drosophila
435	6	2.2	115	16	08EDV4	08EDV4 shewanella
436	6	2.2	115	16	07MMV4	07MMV4 bordetella
437	6	2.2	115	17	07VUV5	07VUV5 bordetella
438	6	2.2	115	17	09V100	09V100 pyrococcus
439	6	2.2	116	10	042353	042353 arabidopsis
440	6	2.2	117	16	091786	091786 pseudomonas
441	6	2.2	117	16	083BY2	083BY2 coxiella bu
442	6	2.2	119	2	084F43	084F43 clostridium
443	6	2.2	120	2	092N12	092N12 staphylococ
444	6	2.2	120	16	087CV5	087CV5 xylella fas
445	6	2.2	120	16	07UPR2	07UPR2 rhodospirill
446	6	2.2	121	16	09JU23	09JU23 neisseria m
447	6	2.2	121	16	081UV8	081UV8 bacillus an
448	6	2.2	122	5	08ST62	08ST62 encephalito
449	6	2.2	122	5	08SC07	08SC07 encephalito
450	6	2.2	122	9	021992	021992 streptococ
451	6	2.2	124	5	09U333	09U333 caenorhabd
452	6	2.2	124	16	09SW84	09SW84 staphylococ
453	6	2.2	125	2	087959	087959 salmonella
454	6	2.2	126	16	08YPY1	08YPY1 anabaena sp
455	6	2.2	126	16	08D5D0	08D5D0 vibrio vuln
456	6	2.2	127	5	08IP12	08IP12 drosophila
457	6	2.2	127	10	09EXR3	09EXR3 nicotiana t
458	6	2.2	127	10	09FXR2	09FXR2 nicotiana t
459	6	2.2	127	12	091GU7	091GU7 epiphyas po
460	6	2.2	127	16	082GR5	082GR5 streptomyc
461	6	2.2	128	2	084E92	084E92 gamma-prote
462	6	2.2	128	2	0848L3	0848L3 gamma-prote
463	6	2.2	128	3	094597	094597 schistosach
464	6	2.2	128	12	089084	089084 variola vir
465	6	2.2	128	16	097IK4	097IK4 clostridium
466	6	2.2	128	16	097CV3	097CV3 clostridium
467	6	2.2	128	16	08U9R3	08U9R3 agrobacteri
468	6	2.2	128	16	086GV2	086GV2 pseudomonas
469	6	2.2	129	11	063911	063911 mus sp. tap
470	6	2.2	129	17	08U116	08U116 pyrococcus
471	6	2.2	130	2	085202	085202 streptococ
472	6	2.2	130	2	08RSP3	08RSP3 streptococ
473	6	2.2	130	2	09F518	09F518 escherichia
474	6	2.2	130	2	09S347	09S347 prevotella
475	6	2.2	130	16	08D8A5	08D8A5 vibrio vuln
476	6	2.2	130	16	087UJ9	087UJ9 pseudomonas
477	6	2.2	130	16	087N42	087N42 vibrio para
478	6	2.2	132	6	0862E8	0862E8 bos taurus
479	6	2.2	132	16	09C7Q3	09C7Q3 pasteurella
480	6	2.2	132	16	08FWR3	08FWR3 bruceella su
481	6	2.2	132	16	0821G2	0821G2 chlamydomon
482	6	2.2	133	2	07WWE6	07WWE6 yersinia al
483	6	2.2	133	5	021887	021887 caenorhabd
484	6	2.2	133	11	0921B5	0921B5 mus musculu
485	6	2.2	133	16	08FMP5	08FMP5 corynebacte
486	6	2.2	133	16	08DCB9	08DCB9 vibrio vuln
487	6	2.2	134	5	015615	015615 entamoeba h
488	6	2.2	134	16	087TK9	087TK9 vibrio para
489	6	2.2	135	2	09F7U0	09F7U0 helicobacte
490	6	2.2	135	5	09G8Q7	09G8Q7 habesia big
491	6	2.2	135	8	07Y1N1	07Y1N1 celatobact
492	6	2.2	135	8	07Y1A0	07Y1A0 celatobact
493	6	2.2	135	8	07Y196	07Y196 celatobact
494	6	2.2	135	8	07Y195	07Y195 celatobact
495	6	2.2	135	8	07Y193	07Y193 celatobact
496	6	2.2	135	8	07Y192	07Y192 celatobact
497	6	2.2	135	8	07Y191	07Y191 celatobact
498	6	2.2	135	8	07Y190	07Y190 celatobact
499	6	2.2	135	8	07Y189	07Y189 celatobact
500	6	2.2	135	8	07Y188	07Y188 celatobact
501	6	2.2	135	8	07Y186	07Y186 celatobact
502	6	2.2	135	8	07Y185	07Y185 celatobact
503	6	2.2	135	8	07Y183	07Y183 celatobact
504	6	2.2	135	8	07Y182	07Y182 celatobact
505	6	2.2	135	8	07Y181	07Y181 platyoster
506	6	2.2	135	8	07Y180	07Y180 drymaplanet
507	6	2.2	135	8	07Y179	07Y179 celatobact
508	6	2.2	135	8	07Y178	07Y178 celatobact
509	6	2.2	135	10	07Y177	07Y177 celatobact
510	6	2.2	135	16	09K1N2	09K1N2 arabidopsis
511	6	2.2	135	16	09K1N2	09K1N2 neisseria m
512	6	2.2	135	16	093JH3	093JH3 streptomyc
513	6	2.2	135	16	07W1C6	07W1C6 bordetella
514	6	2.2	136	5	09AH93	09AH93 bombyx mori
515	6	2.2	136	5	062434	062434 caenorhabd
516	6	2.2	136	6	09N104	09N104 ovis aries
517	6	2.2	136	10	0940S8	0940S8 rosa hybrid
518	6	2.2	137	8	094WU1	094WU1 peribacteri
519	6	2.2	137	16	007780	007780 mycobacteri
520	6	2.2	137	16	08EGV6	08EGV6 shewanella
521	6	2.2	137	16	087J29	087J29 vibrio para
522	6	2.2	137	16	0831Z3	0831Z3 shigella fl
523	6	2.2	138	5	09S1Y3	09S1Y3 mycobacteri
524	6	2.2	138	16	09SRJ7	09SRJ7 caenorhabd
525	6	2.2	138	16	083B89	083B89 coxiella bu
526	6	2.2	139	10	07X7F8	07X7F8 oxyza sativ
527	6	2.2	139	16	089100	089100 bradyrhizob

528	6	2.2	139	16	Q7V985	Q7V985 prochloroco	601	6	2.2	157	2	Q9KRD5	Q9KRD5 streptomyc
529	6	2.2	140	8	Q8WAC3	Q8WAC3 ligia yltie	602	6	2.2	157	1	Q9J904	Q9J904 mus musculu
530	6	2.2	141	16	Q8PIA4	Q8PIA4 streptococ	603	6	2.2	157	11	Q8QVB5	Q8QVB5 rhizobium I
531	6	2.2	141	16	Q8K788	Q8K788 streptococ	604	6	2.2	157	16	Q98B79	Q98B79 rhizobium m
532	6	2.2	141	16	Q88Y81	Q88Y81 lactobacill	605	6	2.2	157	16	Q92S89	Q92S89 rhizobium m
533	6	2.2	142	2	Q9ZAZ8	Q9ZAZ8 streptomyc	606	6	2.2	157	16	Q89DR1	Q89DR1 bradyrhizob
534	6	2.2	142	8	Q9T250	Q9T250 phycophthor	607	6	2.2	157	16	Q82106	Q82106 streptomyc
535	6	2.2	142	16	Q8XL69	Q8XL69 clostridium	608	6	2.2	158	2	Q84EB9	Q84EB9 gamma-prote
536	6	2.2	142	16	Q7UJ87	Q7UJ87 rhodospirell	609	6	2.2	158	4	Q8IX95	Q8IX95 homo sapien
537	6	2.2	143	9	Q9XJDO	Q9XJDO streptococ	610	6	2.2	158	4	Q8IX95	Q8IX95 homo sapien
538	6	2.2	143	16	Q92ZC1	Q92ZC1 rhizobium m	611	6	2.2	158	13	Q9PVJ3	Q9PVJ3 typhlopori
539	6	2.2	143	16	Q8BDJ2	Q8BDJ2 pseudomonas	612	6	2.2	158	16	Q9PFM7	Q9PFM7 xylella fas
540	6	2.2	143	17	Q8PVI4	Q8PVI4 methanosa	613	6	2.2	158	16	Q9I1B5	Q9I1B5 pseudomonas
541	6	2.2	144	2	Q9LAI2	Q9LAI2 pasteurella	614	6	2.2	158	16	Q8XDC7	Q8XDC7 escherichia
542	6	2.2	144	16	Q9CPF9	Q9CPF9 pasteurella	615	6	2.2	158	16	Q8NLD9	Q8NLD9 escherichia
543	6	2.2	144	16	Q8RRM0	Q8RRM0 fusobacteri	616	6	2.2	158	16	Q8AB89	Q8AB89 corynebacte
544	6	2.2	144	16	Q8RGE8	Q8RGE8 fusobacteri	617	6	2.2	158	16	Q8AB89	Q8AB89 corynebacte
545	6	2.2	144	16	Q8BED4	Q8BED4 pseudomonas	618	6	2.2	158	16	Q8TBD0	Q8TBD0 xylella fas
546	6	2.2	144	16	Q8XG86	Q8XG86 salmone	619	6	2.2	159	9	Q858Y7	Q858Y7 bacterioph
547	6	2.2	144	16	Q83R76	Q83R76 shigella fl	620	6	2.2	159	12	P87547	P87547 babanki vir
548	6	2.2	144	17	Q58142	Q58142 pyrococcus	621	6	2.2	159	16	Q8RH59	Q8RH59 listeria mo
549	6	2.2	145	16	Q8PR60	Q8PR60 xanthomonas	622	6	2.2	159	16	Q8RH59	Q8RH59 listeria mo
550	6	2.2	145	16	Q8PW79	Q8PW79 xanthomonas	623	6	2.2	159	16	Q8ZCJ5	Q8ZCJ5 yersinia pe
551	6	2.2	146	6	Q8MJ96	Q8MJ96 equus cabal	624	6	2.2	159	16	Q8INQ7	Q8INQ7 bacillus an
552	6	2.2	146	5	Q9VD97	Q9VD97 dirosophila	625	6	2.2	159	16	Q7UPB1	Q7UPB1 rhodospirell
553	6	2.2	147	5	Q26505	Q26505 scapharca 1	626	6	2.2	160	5	Q8WS63	Q8WS63 glosina mo
554	6	2.2	147	6	Q28523	Q28523 macaca mula	627	6	2.2	160	11	Q8C7Y8	Q8C7Y8 mus musculu
555	6	2.2	147	16	Q8ZNU2	Q8ZNU2 salmone	628	6	2.2	160	16	Q8PJ32	Q8PJ32 xanthomonas
556	6	2.2	147	16	Q89P88	Q89P88 bradyrhizob	629	6	2.2	160	16	Q8PR84	Q8PR84 leptospira
557	6	2.2	147	16	Q7WYX6	Q7WYX6 bordetella	630	6	2.2	161	2	Q8S011	Q8S011 bordetella bu
558	6	2.2	147	16	Q7VX30	Q7VX30 bordetella	631	6	2.2	161	8	Q8HGV9	Q8HGV9 typhlatya m
559	6	2.2	147	16	Q7VBE6	Q7VBE6 prochloroco	632	6	2.2	161	8	Q8HGV8	Q8HGV8 typhlatya m
560	6	2.2	148	1	Q93698	Q93698 sulfolobac	633	6	2.2	161	8	Q8HGV7	Q8HGV7 typhlatya m
561	6	2.2	148	2	Q9RB19	Q9RB19 acinetobact	634	6	2.2	161	8	Q8HFT4	Q8HFT4 typhlatya p
562	6	2.2	148	16	Q9A2M3	Q9A2M3 caulopacter	635	6	2.2	161	8	Q8HFT3	Q8HFT3 typhlatya p
563	6	2.2	148	16	Q97ORA4	Q97ORA4 streptococ	636	6	2.2	161	8	Q8HFT2	Q8HFT2 typhlatya p
564	6	2.2	148	16	Q8BDF1	Q8BDF1 pseudomonas	637	6	2.2	161	8	Q8HBA5	Q8HBA5 typhlatya s
565	6	2.2	148	16	Q87YK6	Q87YK6 pseudomonas	638	6	2.2	161	8	Q8HBS5	Q8HBS5 typhlatya m
566	6	2.2	148	16	Q87YK6	Q87YK6 pseudomonas	639	6	2.2	161	8	Q8HBS5	Q8HBS5 typhlatya m
567	6	2.2	149	2	Q84HK2	Q84HK2 yersinia fas	640	6	2.2	161	16	Q86G55	Q86G55 pseudomonas
568	6	2.2	149	2	Q7WME7	Q7WME7 yersinia al	641	6	2.2	161	16	Q83JH6	Q83JH6 enterococcu
569	6	2.2	149	15	Q7SSW7	Q7SSW7 chimpanzee	642	6	2.2	162	2	Q9PDE8	Q9PDE8 uncultured
570	6	2.2	149	16	Q8YV14	Q8YV14 bruceella me	643	6	2.2	162	12	P87710	P87710 rotavirus a
571	6	2.2	149	16	Q8XC10	Q8XC10 escherichia	644	6	2.2	162	12	Q02476	Q02476 poa semilat
572	6	2.2	149	16	Q8XUL8	Q8XUL8 ralsstonia s	645	6	2.2	162	16	Q8PGC5	Q8PGC5 xanthomonas
573	6	2.2	149	16	Q8XUL8	Q8XUL8 ralsstonia s	646	6	2.2	163	2	Q93IH5	Q93IH5 wolstonia s
574	6	2.2	149	16	Q8PH48	Q8PH48 xanthomonas	647	6	2.2	163	5	Q17155	Q17155 barbatia 11
575	6	2.2	149	16	Q8FOT2	Q8FOT2 leptospira	648	6	2.2	163	16	Q7V0X2	Q7V0X2 prochloroco
576	6	2.2	149	17	Q8PX90	Q8PX90 methanosa	649	6	2.2	164	12	Q9EMW7	Q9EMW7 amsacta moo
577	6	2.2	150	2	Q84EAO	Q84EAO gamma-proce	650	6	2.2	164	17	Q8ZT11	Q8ZT11 pyrobaculum
578	6	2.2	150	16	Q981Z3	Q981Z3 rhizobium 1	651	6	2.2	165	5	Q868X8	Q868X8 drosophila
579	6	2.2	151	2	Q9AHX4	Q9AHX4 carsonella	652	6	2.2	165	16	Q7USL9	Q7USL9 rhodospirell
580	6	2.2	151	16	Q8ABX8	Q8ABX8 bacteroides	653	6	2.2	166	5	Q9VWW7	Q9VWW7 thermoaer
581	6	2.2	151	10	Q89NU0	Q89NU0 bradyrhizob	654	6	2.2	166	16	Q8R7A3	Q8R7A3 chlorophila
582	6	2.2	152	10	Q7XSU0	Q7XSU0 cryza sativ	655	6	2.2	166	16	Q8KCT4	Q8KCT4 thermoanaer
583	6	2.2	152	16	Q97QO98	Q97QO98 streptococ	656	6	2.2	166	16	Q8KCT4	Q8KCT4 chlorophila
584	6	2.2	152	16	Q8WYF4	Q8WYF4 streptococ	657	6	2.2	167	4	Q9P0R0	Q9P0R0 homo sapien
585	6	2.2	153	16	Q92S17	Q92S17 rhizobium m	658	6	2.2	167	4	Q51583	Q51583 homo sapien
586	6	2.2	153	16	Q9ADV5	Q9ADV5 streptomyc	659	6	2.2	167	5	Q77278	Q77278 drosophila
587	6	2.2	153	16	Q92VL2	Q92VL2 streptomyc	660	6	2.2	167	10	Q84KX8	Q84KX8 cucumis mel
588	6	2.2	154	5	Q86MP4	Q86MP4 caenorhabdi	661	6	2.2	167	11	Q8BHQ9	Q8BHQ9 mus musculu
589	6	2.2	154	16	Q82MZ7	Q82MZ7 streptomyc	662	6	2.2	167	16	Q7WNO4	Q7WNO4 bordetella
590	6	2.2	154	16	Q7UZR3	Q7UZR3 prochloroco	663	6	2.2	167	16	Q7WBI2	Q7WBI2 bordetella
591	6	2.2	155	5	Q21024	Q21024 caenorhabdi	664	6	2.2	168	2	Q8GCE0	Q8GCE0 escherichia
592	6	2.2	155	16	Q9S9S7	Q9S9S7 rhizobium 1	665	6	2.2	168	4	Q9UFW8	Q9UFW8 homo sapien
593	6	2.2	156	2	Q9AH07	Q9AH07 rhodococcus	666	6	2.2	168	10	Q8H7F3	Q8H7F3 ataxidipos
594	6	2.2	156	3	Q05660	Q05660 saccharomyc	667	6	2.2	168	16	Q97SW0	Q97SW0 streptococ
595	6	2.2	156	13	Q9PVJ6	Q9PVJ6 eumeces ine	668	6	2.2	168	16	Q8X9P8	Q8X9P8 escherichia
596	6	2.2	156	16	Q8XES7	Q8XES7 escherichia	669	6	2.2	168	16	Q8YRT5	Q8YRT5 bruceella me
597	6	2.2	156	16	Q8RZ27	Q8RZ27 bruceella su	670	6	2.2	168	16	Q8FHP3	Q8FHP3 escherichia
598	6	2.2	156	16	Q8CWV5	Q8CWV5 escherichia	671	6	2.2	168	16	Q8CZ86	Q8CZ86 streptococ
599	6	2.2	156	16	Q87U00	Q87U00 sulfolobus	672	6	2.2	168	16	Q7UCM9	Q7UCM9 shigella fl
600	6	2.2	157	2	Q847J0	Q847J0 pseudomonas	673	6	2.2	168	17	Q9HNFS	Q9HNFS halobacteri

674	6	2.2	168	17	Q8TUP3	Q8TUP3 methanosarc	747	6	2.2	181	16	Q813P1	Q813P1 bacillus ce
675	6	2.2	169	2	Q9AP50	Q9AP50 streptococc	748	6	2.2	181	16	Q7W6G4	Q7W6G4 bordetella
676	6	2.2	169	2	Q9K3A8	Q9K3A8 enterococcu	749	6	2.2	181	16	Q7W4Y6	Q7W4Y6 bordetella
677	6	2.2	169	2	Q84E97	Q84E97 gamma-prote	750	6	2.2	182	10	Q8S0E3	Q8S0E3 cryza sativ
678	6	2.2	169	2	Q7X0P0	Q7X0P0 uncultured	751	6	2.2	182	16	Q8P3E8	Q8P3E8 xanthomonas
679	6	2.2	169	2	Q7WXP5	Q7WXP5 uncultured	752	6	2.2	182	16	Q8NUE8	Q8NUE8 staphylococ
680	6	2.2	169	2	Q7WXP3	Q7WXP3 uncultured	753	6	2.2	183	5	Q26747	Q26747 trypanosoma
681	6	2.2	169	2	Q7WXP2	Q7WXP2 uncultured	754	6	2.2	183	15	Q83C93	Q83C93 coxiella bu
682	6	2.2	169	2	Q7WXP2	Q7WXP2 uncultured	755	6	2.2	184	15	Q8DCX9	Q8DCX9 vibrio vuln
683	6	2.2	169	2	Q7WYN9	Q7WYN9 uncultured	756	6	2.2	184	17	Q8TSV7	Q8TSV7 methanosarc
684	6	2.2	169	2	Q7WYN8	Q7WYN8 uncultured	757	6	2.2	185	2	Q93L86	Q93L86 burkholderi
685	6	2.2	169	2	Q7WYN6	Q7WYN6 uncultured	758	6	2.2	185	16	Q9PD04	Q9PD04 xylella fas
686	6	2.2	169	2	Q7WYN5	Q7WYN5 uncultured	759	6	2.2	185	16	Q8PFM4	Q8PFM4 xanthomonas
687	6	2.2	169	2	Q7WYN4	Q7WYN4 uncultured	760	6	2.2	185	16	Q83E25	Q83E25 coxiella bu
688	6	2.2	169	2	Q7WYN1	Q7WYN1 uncultured	761	6	2.2	186	2	Q9AP03	Q9AP03 burkholderi
689	6	2.2	169	2	Q9BMJ9	Q9BMJ9 homo sapien	762	6	2.2	186	15	Q8NMS8	Q8NMS8 corynebacte
690	6	2.2	169	2	Q9BMD9	Q9BMD9 homo sapien	763	6	2.2	186	16	Q8DE86	Q8DE86 vibrio vuln
691	6	2.2	169	10	Q7Y079	Q7Y079 lactococcu	764	6	2.2	187	2	Q84805	Q84805 streptococc
692	6	2.2	170	8	Q8M1H6	Q8M1H6 chaetosphae	765	6	2.2	187	16	Q7VXK3	Q7VXK3 haemophilus
693	6	2.2	170	9	Q8SC15	Q8SC15 pseudomonas	766	6	2.2	187	16	Q7VXK3	Q7VXK3 anabaena sp
694	6	2.2	170	10	Q94L76	Q94L76 oryza sativ	767	6	2.2	188	11	Q921P1	Q921P1 mus musculu
695	6	2.2	170	10	Q7XCH7	Q7XCH7 oryza sativ	768	6	2.2	189	15	Q8DWN5	Q8DWN5 streptococc
696	6	2.2	170	16	Q9RUD0	Q9RUD0 delnoccocus	769	6	2.2	190	8	Q7YLP0	Q7YLP0 thuja plica
697	6	2.2	170	16	Q97PY6	Q97PY6 streptococc	770	6	2.2	190	17	Q97Y51	Q97Y51 sulfolobus
698	6	2.2	170	16	Q8CYL2	Q8CYL2 streptococc	771	6	2.2	191	5	Q812E8	Q812E8 caenorhadi
699	6	2.2	170	16	Q83J95	Q83J95 shigella fl	772	6	2.2	191	16	Q899R5	Q899R5 ciostriidum
700	6	2.2	171	2	Q523S5	Q523S5 mycoplasma	773	6	2.2	192	15	Q9HW07	Q9HW07 pseudomonas
701	6	2.2	171	16	P73536	P73536 synechocyst	774	6	2.2	193	5	Q8N0F6	Q8N0F6 plasmodium
702	6	2.2	171	16	Q8CUC3	Q8CUC3 staphylococ	775	6	2.2	193	10	Q851P6	Q851P6 oryza sativ
703	6	2.2	171	16	Q81RK1	Q81RK1 bacillus an	776	6	2.2	193	15	Q9DPY2	Q9DPY2 human immun
704	6	2.2	173	16	Q9J5J2	Q9J5J2 chlamydia p	777	6	2.2	193	16	Q82GQ3	Q82GQ3 streptomyc
705	6	2.2	173	16	Q87WN6	Q87WN6 pseudomonas	778	6	2.2	193	16	Q7XK72	Q7XK72 mycobacteri
706	6	2.2	173	17	Q8TNN8	Q8TNN8 methanosarc	779	6	2.2	194	8	Q959M1	Q959M1 ligia occid
707	6	2.2	174	2	Q8KSV6	Q8KSV6 methylocald	780	6	2.2	194	16	Q9RUX7	Q9RUX7 delnoccocus
708	6	2.2	174	2	Q8KSV7	Q8KSV7 methylocald	781	6	2.2	195	4	Q96H61	Q96H61 homo sapien
709	6	2.2	174	2	Q8KT04	Q8KT04 methylocald	782	6	2.2	195	16	Q9RZH6	Q9RZH6 delnoccocus
710	6	2.2	174	2	Q8KT03	Q8KT03 methylocald	783	6	2.2	195	16	Q8KCS5	Q8KCS5 chlorobium
711	6	2.2	174	6	Q97625	Q97625 bos taurus	784	6	2.2	196	10	Q7XKX3	Q7XKX3 cryza sativ
712	6	2.2	175	10	Q9M8J6	Q9M8J6 arabidopsis	785	6	2.2	196	16	Q8U7Q6	Q8U7Q6 agrobacteri
713	6	2.2	175	16	Q8PR21	Q8PR21 xanthomonas	786	6	2.2	196	16	Q88LV1	Q88LV1 pseudomonas
714	6	2.2	176	2	Q544J3	Q544J3 methylocald	787	6	2.2	196	16	Q7W9S2	Q7W9S2 bordetella
715	6	2.2	176	2	Q544I1	Q544I1 methylocald	788	6	2.2	197	5	Q95U01	Q95U01 dictyostelle
716	6	2.2	176	2	Q544I2	Q544I2 methylocald	789	6	2.2	197	15	Q9F3J9	Q9F3J9 streptomyc
717	6	2.2	176	4	Q96BA4	Q96BA4 homo sapien	790	6	2.2	197	17	Q87A17	Q87A17 thermoplasma
718	6	2.2	176	16	Q7VLM3	Q7VLM3 haemophilus	791	6	2.2	197	17	Q97Y39	Q97Y39 plasmodium
719	6	2.2	176	17	Q97AN4	Q97AN4 thermoplasma	792	6	2.2	198	5	Q77005	Q77005 plasmodium
720	6	2.2	177	2	Q9ZH26	Q9ZH26 pseudomonas	793	6	2.2	198	8	Q85M15	Q85M15 paraclobacta
721	6	2.2	177	2	Q9S6C0	Q9S6C0 pseudomonas	794	6	2.2	198	11	Q8R4B9	Q8R4B9 mesocricetu
722	6	2.2	177	15	Q8JF06	Q8JF06 human immun	795	6	2.2	198	16	Q8ZLE8	Q8ZLE8 salmonella
723	6	2.2	177	16	Q8ZQJ8	Q8ZQJ8 streptomyc	796	6	2.2	198	16	Q8XKRO	Q8XKRO escherichia
724	6	2.2	178	15	P88235	P88235 human immun	797	6	2.2	198	16	Q8Z2S2	Q8Z2S2 salmonella
725	6	2.2	178	16	Q25673	Q25673 helicobacte	798	6	2.2	198	16	Q8Z2S2	Q8Z2S2 salmonella
726	6	2.2	178	16	Q9ZM28	Q9ZM28 helicobacte	799	6	2.2	198	16	Q8Z2S2	Q8Z2S2 salmonella
727	6	2.2	178	16	Q8BDT8	Q8BDT8 pseudomonas	800	6	2.2	198	16	Q83J84	Q83J84 shigella fl
728	6	2.2	179	4	Q8TC97	Q8TC97 homo sapien	801	6	2.2	198	17	Q9YBK5	Q9YBK5 aeropyrum p
729	6	2.2	179	8	Q8WAC6	Q8WAC6 ligia exoti	802	6	2.2	199	4	Q9HAMS	Q9HAMS homo sapien
730	6	2.2	179	10	Q9LFI8	Q9LFI8 arabidopsis	803	6	2.2	199	11	Q9EPH2	Q9EPH2 rattus norv
731	6	2.2	179	11	Q9D9Y3	Q9D9Y3 mus musculu	804	6	2.2	199	16	Q9CIU5	Q9CIU5 pasteurella
732	6	2.2	179	11	Q9DA64	Q9DA64 mus musculu	805	6	2.2	200	2	Q9ALU8	Q9ALU8 enterobacte
733	6	2.2	179	16	Q8FLE7	Q8FLE7 corynebacte	806	6	2.2	200	5	Q817Z1	Q817Z1 metapneus
734	6	2.2	179	16	Q89YF0	Q89YF0 bradyrhizob	807	6	2.2	200	16	Q9NM52	Q9NM52 thizobium m
735	6	2.2	180	2	Q57025	Q57025 escherichia	808	6	2.2	200	16	Q8EF71	Q8EF71 bruceella su
736	6	2.2	180	3	Q74882	Q74882 schistosach	809	6	2.2	200	16	Q8ZDX3	Q8ZDX3 streptomyc
737	6	2.2	180	10	Q9LGC7	Q9LGC7 oryza sativ	810	6	2.2	201	3	Q8US91	Q8US91 schistosach
738	6	2.2	180	16	Q9KSP8	Q9KSP8 bacillus ha	811	6	2.2	201	6	Q8WNV4	Q8WNV4 ovis aries
739	6	2.2	180	16	Q98QJ6	Q98QJ6 mycoplasma	812	6	2.2	201	8	Q9GID2	Q9GID2 caridina ze
740	6	2.2	180	16	Q83RX9	Q83RX9 shigella fl	813	6	2.2	201	8	Q9G470	Q9G470 caridina ze
741	6	2.2	181	2	Q9XBM7	Q9XBM7 rhodobacter	814	6	2.2	201	8	Q9G1D3	Q9G1D3 caridina ze
742	6	2.2	181	2	Q9ETV6	Q9ETV6 corynebacte	815	6	2.2	201	8	Q9G1D1	Q9G1D1 caridina ze
743	6	2.2	181	4	Q14608	Q14608 homo sapien	816	6	2.2	201	8	Q9G469	Q9G469 caridina ze
744	6	2.2	181	16	Q9BZM4	Q9BZM4 homo sapien	817	6	2.2	201	17	Q9G531	Q9G531 archaeoglob
745	6	2.2	181	16	Q9KPE7	Q9KPE7 vibrio chol	818	6	2.2	201	17	Q9G531	Q9G531 archaeoglob
746	6	2.2	181	16	Q9X826	Q9X826 streptomyc	819	6	2.2	202	2	Q8KVT1	Q8KVT1 staphylococ

820	6	2.2	202	2	Q9KWM2	Q9Kwm2 staphylococ	893	6	2.2	219	5	Q9VJR3	Q9vj3 drosophila
821	6	2.2	202	2	Q53757	Q53757 staphylococ	894	6	2.2	219	5	Q8IDV4	Q8idv4 plasmodium
822	6	2.2	202	5	Q8N0G1	Q8n0g1 plasmodium	895	6	2.2	219	8	Q8WBA2	Q8wb2 bnf1o panthe
823	6	2.2	202	16	Q97LE7	Q97le7 clostridium	896	6	2.2	219	8	Q8W7U5	Q8w7u5 ligia perki
824	6	2.2	202	16	Q8DLE8	Q8dle8 synecchococ	897	6	2.2	219	8	Q8WAB8	Q8wab8 ligia hawai
825	6	2.2	202	16	Q7VGM4	Q7vgm4 helicobacte	898	6	2.2	219	8	Q8WAB5	Q8wab5 ligia perki
826	6	2.2	203	5	Q9UYI2	Q9uly2 caenorhabdi	899	6	2.2	219	8	Q8WAC2	Q8wac2 ligia hawai
827	6	2.2	203	5	Q45289	Q45289 caenorhabdi	900	6	2.2	219	8	Q8WAC1	Q8wac1 ligia hawai
828	6	2.2	203	13	Q90XY0	Q90xy0 fungu rubrip	901	6	2.2	219	10	Q9SP76	Q9spd7 gossypium a
829	6	2.2	203	16	Q8XND6	Q8xnd6 clostridium	902	6	2.2	219	10	Q9S7S9	Q9s7s9 gossypium 1
830	6	2.2	203	16	Q8CUD5	Q8cud5 staphylococ	903	6	2.2	219	16	Q9PBR3	Q9pb3 xylella fas
831	6	2.2	203	16	Q89SRS	Q89ers bradyrhizob	904	6	2.2	219	16	Q7W329	Q7wb29 bordetella
832	6	2.2	203	16	Q83L89	Q83l89 shigella fl	905	6	2.2	220	16	Q9S2X4	Q9s2x4 streptomyce
833	6	2.2	204	16	Q8YIP0	Q8yip0 bruceella me	906	6	2.2	220	16	Q89ZS6	Q89z66 bacteroides
834	6	2.2	205	2	Q8GKW7	Q8gkw7 streptococ	907	6	2.2	221	9	Q21979	Q21979 bacterioph
835	6	2.2	205	16	Q51738	Q51738 borrelia bu	908	6	2.2	221	10	Q8VY90	Q8vy90 arabisdops
836	6	2.2	205	16	Q8BRK3	Q8brk3 pseudomonas	909	6	2.2	221	10	Q8LD04	Q8ld04 arabisdops
837	6	2.2	205	16	Q7U7T8	Q7u7t8 synecchococ	910	6	2.2	221	16	Q8XMR6	Q8xmr6 ralsconia s
838	6	2.2	206	2	Q44801	Q44801 borrelia bu	911	6	2.2	221	17	Q97UD7	Q97ud7 sulfolobus
839	6	2.2	206	8	Q8WAB7	Q8wab7 ligia hawai	912	6	2.2	222	2	P72322	P72322 rhodospirill
840	6	2.2	206	8	Q8WACO	Q8waco ligia hawai	913	6	2.2	223	16	Q8FEW9	Q8f9w9 fusobacteri
841	6	2.2	206	8	Q8WAB9	Q8wab9 ligia hawai	914	6	2.2	223	16	Q8FM15	Q8fms15 corynebacte
842	6	2.2	206	8	Q8W7H2	Q8w7h2 ligia hawai	915	6	2.2	223	16	Q8E3S8	Q8e3s8 streptococ
843	6	2.2	206	16	Q8YER8	Q8yer8 bruceella me	916	6	2.2	223	16	Q8DYC2	Q8dyr2 streptococ
844	6	2.2	206	16	Q8G317	Q8g317 bruceella su	917	6	2.2	223	16	Q8B1Y6	Q8b1y6 lactobacill
845	6	2.2	206	16	Q87SY7	Q87sy7 vibrio para	918	6	2.2	223	16	Q8Z5K5	Q8z5t5 streptomyce
846	6	2.2	207	5	Q8N0G0	Q8n0g0 plasmodium	919	6	2.2	223	16	Q7UEA5	Q7uea5 rhodospirell
847	6	2.2	207	5	Q15675	Q15675 caenorhabdi	920	6	2.2	224	2	Q9AF13	Q9af13 shigella fl
848	6	2.2	207	10	Q91E71	Q91e11 arabidopsi	921	6	2.2	224	2	Q8VTH8	Q8vth8 shigella fl
849	6	2.2	207	16	Q7UTY0	Q7uty0 rhodospirell	922	6	2.2	224	10	Q9FHS3	Q9fhs3 arabisdops
850	6	2.2	208	2	Q849C2	Q849c2 haemophilus	923	6	2.2	224	16	Q9BRT2	Q9brt2 rhizobium 1
851	6	2.2	208	8	Q85MR8	Q85mr8 ephemerella	924	6	2.2	224	16	Q8G2U7	Q8g2u7 bruceella su
852	6	2.2	208	8	Q85M07	Q85mq7 periplaneta	925	6	2.2	224	16	Q81ZQ7	Q81zq7 shigella fl
853	6	2.2	208	8	Q85M19	Q85m19 chrysotoxum	926	6	2.2	224	16	Q9AFM6	Q9afm6 shigella fl
854	6	2.2	208	8	Q85M07	Q85mq7 eurylophell	927	6	2.2	224	16	Q7UD61	Q7ud61 shigella fl
855	6	2.2	208	16	Q8NM00	Q8nmq0 corynebacte	928	6	2.2	225	16	Q8YUS5	Q8yus5 bradyrhizob
856	6	2.2	208	16	Q8BET3	Q8bet3 shewanella	929	6	2.2	226	16	Q8Y271	Q8y271 anabaena sp
857	6	2.2	208	16	Q7WCY6	Q7wcy6 bordetella	930	6	2.2	226	16	Q8BWT1	Q8bwt1 pseudomonas
858	6	2.2	209	2	Q9FBW4	Q9fbw4 bordetella	931	6	2.2	226	16	Q8BWT1	Q8bwt1 pseudomonas
859	6	2.2	209	2	Q21621	Q21621 caenorhabdi	932	6	2.2	227	2	Q8ZPY5	Q8zpy5 streptomyce
860	6	2.2	209	16	Q9HX40	Q9hx40 pseudomonas	933	6	2.2	227	2	Q54768	Q54768 synecchococ
861	6	2.2	209	16	Q8NNZ4	Q8nnz4 corynebacte	934	6	2.2	227	5	Q8KIP0	Q8kip0 rhizobium 1
862	6	2.2	209	16	Q8A5E5	Q8ase5 bacteroides	935	6	2.2	227	5	Q8BDP6	Q8bdp6 sepiia offic
863	6	2.2	210	5	Q966C8	Q966c8 caenorhabdi	936	6	2.2	227	5	Q7WKP5	Q7wkp5 dictyosteli
864	6	2.2	210	16	Q9RTF2	Q9rtf2 deinococcus	937	6	2.2	227	16	Q8ZJ00	Q8zj00 salmonella
865	6	2.2	210	16	Q55649	Q55649 synecchocyt	938	6	2.2	227	16	Q8Z0S8	Q8z0s8 salmonella
866	6	2.2	210	16	Q9RG61	Q9rg61 pseudomonas	939	6	2.2	227	17	Q8TUS3	Q8tus3 methanopyru
867	6	2.2	210	16	Q8F1E7	Q8f1e7 leptospira	940	6	2.2	228	5	Q81H06	Q81h06 drosophila
868	6	2.2	210	16	Q7UID6	Q7uid6 rhodospirell	941	6	2.2	228	10	Q8LSU4	Q8lsu4 arabisdops
869	6	2.2	211	5	Q81RH2	Q81rh2 drosophila	942	6	2.2	228	10	Q9FOA1	Q9foa1 zea mays (m
870	6	2.2	211	16	Q8PBC2	Q8pbc2 xanthomonas	943	6	2.2	228	10	Q9M1R7	Q9m1r7 arabisdops
871	6	2.2	211	16	Q896B8	Q896b8 clostridium	944	6	2.2	228	16	Q537E0	Q537e0 mycobacteri
872	6	2.2	213	10	Q8VY51	Q8vy51 arabisdops	945	6	2.2	228	16	Q87PW8	Q87pw8 vibrio para
873	6	2.2	213	16	Q99YN3	Q99yn3 streptococ	946	6	2.2	228	17	Q7UIQ2	Q7uiq2 mycobacteri
874	6	2.2	213	16	Q99W74	Q99w74 staphylococ	947	6	2.2	228	17	P94946	P94946 methanopyru
875	6	2.2	213	16	Q8P008	Q8p008 streptococ	948	6	2.2	230	13	Q7SYJ9	Q7syj9 brachydanio
876	6	2.2	213	16	Q8PPF9	Q8ppf9 streptococ	949	6	2.2	230	16	Q8NP14	Q8np14 corynebacte
877	6	2.2	214	16	Q92W95	Q92w95 rhizobium m	950	6	2.2	230	16	Q8BQC0	Q8bqc0 pseudomonas
878	6	2.2	214	16	Q8E7J5	Q8e7j5 streptococ	951	6	2.2	231	2	Q45961	Q45961 clostridium
879	6	2.2	214	16	Q8E1L9	Q8e1l9 streptococ	952	6	2.2	231	2	Q69113	Q69113 burkholderi
880	6	2.2	215	6	Q81031	Q81031 bos taurus	953	6	2.2	231	11	Q8CF79	Q8cf79 mus musculu
881	6	2.2	215	10	Q9MJU5	Q9mju5 nicotiana t	954	6	2.2	231	16	Q88CB7	Q88cb7 pseudomonas
882	6	2.2	215	16	Q879E8	Q879e8 streptococ	955	6	2.2	231	17	Q9Y9D7	Q9y9d7 aeropyrum p
883	6	2.2	215	17	Q8PW73	Q8pw73 methanosarc	956	6	2.2	231	17	Q978V6	Q978v6 thermoplasm
884	6	2.2	216	2	Q8KTY4	Q8kty4 uncultured	957	6	2.2	232	2	Q7WVW4	Q7wvw4 alcaligenes
885	6	2.2	216	2	Q9F243	Q9f243 actinobacill	958	6	2.2	232	2	Q9YUIC1	Q9yui1 drosophila
886	6	2.2	216	16	Q7WVW5	Q7wvj5 bordetella	959	6	2.2	232	16	Q8XW02	Q8xw02 ralsconia s
887	6	2.2	216	16	Q7VZAT	Q7vzat bordetella	960	6	2.2	232	16	Q8DTB6	Q8dtb6 streptococ
888	6	2.2	217	8	Q8WU66	Q8wue6 pteris hiru	961	6	2.2	232	16	Q7UC12	Q7uc12 shigella fl
889	6	2.2	217	16	Q8DQL8	Q8ddl8 streptococ	962	6	2.2	233	16	Q9AMJ90	Q9aj90 actinobacill
890	6	2.2	218	8	Q961J3	Q961j3 homo sapien	963	6	2.2	233	16	Q8XZN3	Q8xzn3 ralsconia s
891	6	2.2	218	8	Q8WAB6	Q8wab6 ligia perki	964	6	2.2	233	16	Q8DSX2	Q8dsx2 vibrio vuln
892	6	2.2	218	17	Q8THJ3	Q8thj3 methanosarc	965	6	2.2	234	16	Q9BLP6	Q9blp6 rhizobium 1

```

966 6 2.2 235 16 Q9CJY2 Q9CJY2 pasteurella
967 6 2.2 236 2 Q8KMA1 Q8KMA1 ruegeria sp
968 6 2.2 236 12 Q8Q575 Q8Q575 chimpanzee
969 6 2.2 236 16 Q8NMM0 Q8NMM0 corynebace
970 6 2.2 236 16 Q8AL65 Q8AL65 pseudomons
971 6 2.2 237 2 Q9AMC6 Q9AMC6 streptococ
972 6 2.2 237 2 Q8KSP9 Q8KSP9 streptomyce
973 6 2.2 237 2 Q848N0 Q848N0 gamma-prote
974 6 2.2 237 2 Q848M0 Q848M0 gamma-prote
975 6 2.2 237 16 Q9KRM4 Q9KRM4 vibrio chol
976 6 2.2 237 16 Q97RU0 Q97RU0 streptococc
977 6 2.2 237 16 Q8CZ07 Q8CZ07 streptococc
978 6 2.2 238 2 Q88988 Q88988 mycoplasma
979 6 2.2 238 2 Q8KXZ0 Q8KXZ0 desulfobacu
980 6 2.2 238 16 Q89PX3 Q89PX3 bradyrhizob
981 6 2.2 239 10 Q9LFJ3 Q9LFJ3 arabidopsis
982 6 2.2 239 16 Q92NZ4 Q92NZ4 rhizobium m
983 6 2.2 239 16 Q8UDK5 Q8UDK5 agrobacteri
984 6 2.2 239 17 Q97WM0 Q97WM0 sulfolobus
985 6 2.2 239 17 Q8TPX2 Q8TPX2 methanobac
986 6 2.2 239 17 Q8PTU3 Q8PTU3 methanobac
987 6 2.2 240 2 Q93LU0 Q93LU0 streptococ
988 6 2.2 240 11 Q99UX2 Q99UX2 mus musculi
989 6 2.2 240 16 Q8XBD7 Q8XBD7 escherichia
990 6 2.2 240 16 Q8FF80 Q8FF80 escherichia
991 6 2.2 240 16 Q83K55 Q83K55 shigella fl
992 6 2.2 241 5 Q9UYO0 Q9UYO0 leishmania
993 6 2.2 241 16 Q8ZSN1 Q8ZSN1 salmonella
994 6 2.2 241 16 Q8ZS07 Q8ZS07 nitrosomona
995 6 2.2 242 3 Q9P8A6 Q9P8A6 agarticus bi
996 6 2.2 242 10 Q9F155 Q9F155 arabidopsis
997 6 2.2 242 16 Q92FD0 Q92FD0 listeria in
998 6 2.2 242 16 Q8YAJ6 Q8YAJ6 listeria mo
999 6 2.2 242 16 Q9S1U9 Q9S1U9 streptomyce
1000 6 2.2 242 16 Q8BPF0 Q8BPF0 pseudomons

```

ALIGNMENTS

RESULT 1

```

ID Q7WM93 PRELIMINARY; PRT; 270 AA.
AC Q7WM93;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 28kDa outer membrane lipoprotein.
GN HUPA OR HD1100.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3500HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017153; AAP95966.1; -.
KW Lipoprotein, Complete proteome.
SQ SEQUENCE 270 AA; 29535 MW; FC077F10B13304AD CRC64;

```

Query Match 5.8%; Score 16; DB 16; Length 270;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 100 KDSEKGLNLYVGN 115
 DB 95 KDSEKGLNLYVGN 110

```

RESULT 2
ID Q83MC6 PRELIMINARY; PRT; 271 AA.
AC Q83MC6;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative lipoprotein.
GN YABC OR SF0188 OR S0190.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RC MEDLINE=22590774; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / Serotype 2a;
RC MEDLINE=22590774; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
RN [3]
RC EMBL; AE015055; AAN41850.1; -.
DR EMBL; AE016978; AAP15730.1; -.
DR InterPro: IPR004872; Lipoprotein 9.
DR InterPro: IPR000437; Prok_lipoprot_S.
DR InterPro: IPR004478; Yabc.
DR Pfam: PF03180; Lipoprotein 9; 1.
DR TIGRFAMs: TIGR00363; TIGR00363; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 271 AA; 29417 MW; C3D7D01F85CD2BFE CRC64;

```

Query Match 4.0%; Score 11; DB 16; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 224 FVEDKDSPPVN 234
 DB 219 FVEDKDSPPVN 229

RESULT 3

```

ID Q7WIF5 PRELIMINARY; PRT; 271 AA.
AC Q7WIF5;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative exported protein.
GN BB2896.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RC MEDLINE=22827954; PubMed=12910271;

```


RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640445; CAE33388.1; -.
 KW Complete proteome.
 SQ SEQUENCE 271 AA; 29643 MW; FC5BEC19A837AB06 CRC64;

Query Match 3.6%; Score 10; DB 16; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GATIAVPNDP 146
 DB 127 GATIAVPNDP 136

RESULT 4
 ID Q7VM6 PRELIMINARY; PRT; 271 AA.

DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP1292.

OS *Bordetella pertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; *Bordetella*.
 OX NCBI_TaxID=520;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-569 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640445; CAE33388.1; -.
 KW Complete proteome.
 SQ SEQUENCE 271 AA; 29614 MW; E75BEC1C2337AB06 CRC64;

Query Match 3.6%; Score 10; DB 16; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GATIAVPNDP 146
 DB 127 GATIAVPNDP 136

RESULT 5
 ID Q7W6I3 PRELIMINARY; PRT; 281 AA.
 AC Q7W6I3;

DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP2926.
 OS *Bordetella parapertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; *Bordetella*.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-567;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Quail M.A.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640432; CAE38219.1; -.
 KW Complete proteome.
 SQ SEQUENCE 281 AA; 30591 MW; C97FA51E40E3315 CRC64;

Query Match 3.6%; Score 10; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GATIAVPNDP 146
 DB 137 GATIAVPNDP 146

RESULT 6
 ID O69441 PRELIMINARY; PRT; 259 AA.
 AC O69441;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE 29 kDa immunogenic protein.
 OS *Legionella pneumophila*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 OC Legionellaceae; *Legionella*.
 OX NCBI_TaxID=446;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Corby;
 RA Stendel C., Helbig J., Lueck C.,
 RT "Cloning and antigenic characterization of a 29 kDa protein from
 RT *Legionella pneumophila* strain Corby.";
 RL Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: A005668; CA06664.1; -.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok_lipoprote_S.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PFO3180; Lipoprotein_9; 1.
 DR TIGRFAMs: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 SQ SEQUENCE 259 AA; 28614 MW; 17FDB7AF6C4ADA66 CRC64;

Query Match 3.3%; Score 9; DB 2; Length 259;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVNDPSN 148
 ID IAVNDPSN 148

Db 125 IAVNPDPN 133

RESULT 7

Q7WEM2 PRELIMINARY; PRT; 262 AA.

AC Q7WEM2; 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DE Putative exported protein.

GN BB4519.

OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=518;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RB50 / ATCC BAA-588;

RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebald M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,

RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Stevens K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Umwin L., Whitehead S., Barrrell B.G., Maskell D.J.;

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica.";

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640450; CAB34882.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28490 MW; 0DC80B1739FF4663 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RESULT 8

Q7W317 PRELIMINARY; PRT; 262 AA.

AC Q7W317; 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DE Putative exported protein.

GN BP4046.

OS Bordetella parapertussis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=519;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=12822 / ATCC BAA-587;

RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebald M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,

RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Stevens K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Umwin L., Whitehead S., Barrrell B.G., Maskell D.J.;

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica.";

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640412; CAB44716.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28524 MW; E0A56B1A59FF4EC9 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica.";

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640435; CAB39329.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28506 MW; D4C18652307EBB2A CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RESULT 9

Q7VSM1 PRELIMINARY; PRT; 262 AA.

AC Q7VSM1; 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DE Putative exported protein.

GN BP0385.

OS Bordetella pertussis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=520;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;

RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebald M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,

RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,

RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Stevens K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Umwin L., Whitehead S., Barrrell B.G., Maskell D.J.;

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica.";

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640412; CAB44716.1; -

KM Complete proteome.

SQ SEQUENCE 262 AA; 28524 MW; E0A56B1A59FF4EC9 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 262;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 ELKDGATTA 141

Db 119 ELKDGATTA 127

RESULT 10

O8YA74 PRELIMINARY; PRT; 273 AA.

AC O8YA74; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DE Putative lipoprotein.

GN LMO0285.

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OC NCBI_TaxID=1639;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Claeser P., Frangeul L., Buchrieser C., Ruenick C., Amend A.,
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
 RA Charbit A., Checourant F., Couve E., de Daruvar A., Deloux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Eutican K.-D., Fehli H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkac G.,
 RA Medueno E., Maitournan A., Mata Vicente U., Ng E., Negari H.,
 RA Nordstrek G., Novella S., de Pablos B., Perez-Diaz U.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland U.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of *Listeria species*.";
 RL Science 294:849-852(2001).
 RM EMBL: A1591974; CAD00812.1; -
 DR PIR: A1110; A1110.
 DR Listlist; LMO00285; -
 DR InterPro: IPR004872; Lipoprotein 9.
 DR InterPro: IPR004437; Prok_lipoprot_S.
 DR InterPro: IPR004478; YaeC.
 DR Pfam: PF03180; Lipoprotein 9; 1.
 DR TIGRPFAM: TIGR00363; TIGR00363; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Lipoprotein; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 273 AA; 29811 MW; AC874D27BBA3AABC CRC64;

Query Match 3.3%; Score 9; DB 16; Length 273;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVNPDSN 148
 Db 136 IAVNPDSN 144

RESULT 11
 Q7WML8 PRELIMINARY; PRT; 276 AA.
 AC Q7WML8;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative transglycosylase.
 GN B1373.
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Rabinowitsch E., Rutter S., Norbertczak H., O'Neill S., Ormond D., Price C.,
 RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Umwin L., Whitehead S., Barrall B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640441; CAB31871.1; -
 KM Complete proteome.
 SQ SEQUENCE 276 AA; 28845 MW; C42904282B560533 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 276;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALASGIALA 19
 Db 266 ALASGIALA 274

RESULT 12
 Q7WBS1 PRELIMINARY; PRT; 276 AA.
 AC Q7WBS1;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative transglycosylase.
 GN BPI157.
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Umwin L., Whitehead S., Barrall B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640426; CAB36458.1; -
 KM Complete proteome.
 SQ SEQUENCE 276 AA; 28845 MW; C42904282B560533 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 276;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALASGIALA 19
 Db 266 ALASGIALA 274

RESULT 13
 Q7VZ71 PRELIMINARY; PRT; 276 AA.
 AC Q7VZ71;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative transglycosylase.
 GN BPI061.
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Robama 1 / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,

RA Rabbihowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton U., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640414; CAB41350.1; -.
 KW Complete proteome.
 SQ SEQUENCE 276 AA; 28879 MW; C0126189C59B362 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 276;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ALASGIALA 19
 |||||
 DB 266 ALASGIALA 274

RESULT 14
 Q8YD39 PRELIMINARY; PRT; 278 AA.
 AC Q8YD39;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN BMEI1038.
 OS *Brucella melitensis*.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Brucellaceae; Brucella.
 CX NCBI_Taxid=29459;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=16W / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujar C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselcorn R., Kyriades N., Overbeek R.,
 RT "The genome sequence of the facultative intracellular pathogen
 RT *Brucella melitensis*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL; AEO09671; AAL53580.1; -.
 DR PIR; A13551; A13551.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR005311; Tat.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR01409; TAT signal seq; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 KW Complete proteome.

SQ SEQUENCE 278 AA; 30606 MW; C0126189C59B362 CRC64;
 Query Match 3.3%; Score 9; DB 16; Length 278;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 176 DIVENPKKL 184
 |||||
 DB 177 DIVENPKKL 185

RESULT 15
 Q8FV86 PRELIMINARY; PRT; 278 AA.
 AC Q8FV86;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Lipoprotein, YaeC family.
 GN BRA0960.

OS *Brucella suis*.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 CX NCBI_Taxid=29461;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=1338 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seehardt R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beaman M.J.,
 RA Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Hailing S.M., Boyle S.M., Fraser C.M.,
 RT "The *Brucella suis* genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AB014589; AAN34131.1; -.
 DR PIR; A13551; A13551.
 DR TIGR; BRA0960; -.

DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR005311; Tat.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR01409; TAT signal seq; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 KW Complete proteome.
 SQ SEQUENCE 278 AA; 30606 MW; C0126189C59B362 CRC64;

Query Match 3.3%; Score 9; DB 16; Length 278;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 176 DIVENPKKL 184
 |||||
 DB 177 DIVENPKKL 185

RESULT 16
 Q9UYT7 PRELIMINARY; PRT; 89 AA.
 AC Q9UYT7;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Molycoprotein converting factor, subunit 1 (MOAD).
 GN PYRAH14200 OR PAB3357.
 OS *Pyrococcus abyssi*.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;
 OC Pyrococcus.

OC NCBI_Taxid=29292;
 RP SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RA Helig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 RT structure and evolution,"
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ248287; CAB50325.1; -.
 DR PIR; H75053; H75053.
 DR HSP; P30748; IFMO.
 DR GO; GO:0006790; P:sulfur metabolism; IEX.
 DR InterPro; IPR003749; This.
 DR Pfam; PF02597; This; 1.
 KW Complete proteome.
 SQ SEQUENCE 89 AA; 10252 MW; C937B0E376876E2C CRC64;

Query Match 2.9%; Score 8; DB 17; Length 89;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 ELKGGATT 140
 |||||

Db 23 ELKDGATI 30

RESULT 17

Q9AGI3 PRELIMINARY; PRT; 255 AA.

AC Q9AGI3; (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DE Hypothetical protein.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NC NCB1_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2137446; PubMed=11481492;
 RA Atora S.K., Bangera M., Lory S., Rampal R.;
 RT "A genomic island in Pseudomonas aeruginosa carries the determinants
 of flagellin glycosylation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9342-9347(2001).
 DR EMBL; AF332547; AAK1535.1; -
 DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR InterPro; IPR003329; Cytidylyl_trans.
 DR Pfam; PF02348; CTP_transf_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 255 AA; 28299 MW; 969DD8BA564D53B4 CRC64;

Query Match 2.9%; Score 8; DB 2; Length 255;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 VAEVAGV 58
 DB 80 VAEVAGV 87

RESULT 18

Q7XSM6 PRELIMINARY; PRT; 259 AA.

AC Q7XSM6; (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DE Hypothetical protein.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NC NCB1_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CF5;
 RA Atora S.K., Wolfgang M.C., Lory S., Rampal R.;
 RT "Sequence polymorphism in the glycosylation island and flagellins of
 Pseudomonas aeruginosa.";
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY280453; AAP35723.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 259 AA; 28789 MW; 2B60BD2857EA9038 CRC64;

Query Match 2.9%; Score 8; DB 2; Length 259;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 VAEVAGV 58
 DB 84 VAEVAGV 91

RESULT 19
 Q8XNL7 PRELIMINARY; PRT; 259 AA.
 ID Q8XNL7

AC Q8XNL7;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Probable transcriptional regulator.
 GN CPE0316.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 NC NCB1_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / Type A;
 RX MEDLINE=2164373; PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL; AP003165; BAB0022.1; -
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR01034; HTH_Deor.
 DR Pfam; PF00455; deor; 1.
 DR SMART; SM00420; HTH_Deor; 1.
 DR PROSITE; PS00894; HTH_Deor_FAMILY; 1.
 KW Complete proteome.
 SQ SEQUENCE 259 AA; 29210 MW; 83A04B3A2E6F1A2E CRC64;

Query Match 2.9%; Score 8; DB 16; Length 259;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 LEKQGLIK 163
 DB 43 LEKQGLIK 50

RESULT 20

Q8U7G0 PRELIMINARY; PRT; 259 AA.

AC Q8U7G0;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Outer membrane lipoprotein.
 GN ATU4489 OR AGR_L_761.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NC NCB1_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitejina U.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D., Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kuyavlin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Xu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RN Science 294:2317-2323 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Gurello B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,

RA Houniel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Planagan C., Crowell C., Gerson J., Lomo C., Sear C., Strud G.,
 RA Cleio C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT *Agrobacterium tumefaciens* C58.",
 RT Science 294:2323-2328(2001).
 RL EMBL; AF009377; AA145283.1; -
 DR EMBL; AF008237; AAK88953.1; -
 DR PIR; AE3108; AE3108.
 DR PIR; G98178; G98178.
 DR InterPro; IPR004872; Lipoprotein_9.
 DR InterPro; IPR004478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR Trifam; TIGR00363; TIGR00363; 1.
 KM Lipoprotein, Complete proteome.
 SQ SEQUENCE 259 AA; 27934 MW; F7D12ED13960861A CRC64;

Query Match 2.9%; Score 8; DB 16; Length 259;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 LNEKDA 140
 |||||
 Db 116 LNEKDA 123

RESULT 21
 Q7W4E3 PRELIMINARY; PRT; 265 AA.

AC Q7W4E3 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BBA169
 OS *Bordetella bronchiseptica* (Alcaligenes bronchiseptus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
 RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640449; CAE34533.1; -
 KW Complete proteome.
 SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B566885B CRC64;

Query Match 2.9%; Score 8; DB 16; Length 265;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 LNEKDA 138
 |||||
 Db 120 LNEKDA 127

RESULT 22
 Q7W4E3 PRELIMINARY; PRT; 265 AA.

AC Q7W4E3 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BPP3723.
 OS *Bordetella parapertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
 RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640434; CAE39006.1; -
 KW Complete proteome.
 SQ SEQUENCE 265 AA; 28715 MW; 4CA85474B566885B CRC64;

Query Match 2.9%; Score 8; DB 16; Length 265;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 LNEKDA 138
 |||||
 Db 120 LNEKDA 127

RESULT 23
 Q7W4E3 PRELIMINARY; PRT; 265 AA.

AC Q7W4E3 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative exported protein.
 GN BP2818.
 OS *Bordetella pertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 CC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
 RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640419; CAE33091.1; -
 KW Complete proteome.
 SQ SEQUENCE 265 AA; 28743 MW; 6B585474B5668844 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 265;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

131 LNEKDDGA 138
 |||||
 DB 120 LNEKDDGA 127

RESULT 24
 Q8P4S9 PRELIMINARY; PRT; 266 AA.

AC Q8P4S9;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Outer membrane protein.
 GN XCC3628.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 CX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RC MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria U.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Secubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL: AE012484; AAM42898.1;
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0005215; F:transporter activity; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR004872; Lipoprotein_9.
 DR InterPro: IPR000437; Prok lipoprot_S.
 DR InterPro: IPR000531; TonB box.
 DR Pfam: PF03180; Lipoprotein_9.1.
 DR PROSITE: PS00013; PROXR_LIPOPROTEIN_1.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1.1.
 KW Complete proteome.
 SQ SEQUENCE 266 AA; 28870 MW; 9023F0E86CD3B90 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 266;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

237 VARADND 244
 |||||
 DB 227 VARADND 234

RESULT 25
 Q7V195 PRELIMINARY; PRT; 268 AA.

AC Q7V195;
 DT 01-OCT-2003 (TReMBLrel. 25, Created)

DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Hypothetical protein yaeC.
 GN YAE_C OR H00713.
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 CX NCBI_TaxID=32025;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 51449 / 3B1;
 RC MEDLINE=22709201; PubMed=12810954;
 RA Suerbaum S., Josephans C., Stezenbach T., Drescher B., Brandt P.,
 RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Hollander D.B., Klein K., Koenig J., Macko L., Mendz G.L., Nykatura G.,
 RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.,
 RT "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
 DR EMBL: AE017146; AAP7310.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 268 AA; 29314 MW; 20948E0B2918109F CRC64;

Query Match 2.9%; Score 8; DB 16; Length 268;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

175 LDIYENPK 182
 |||||
 DB 168 LDIYENPK 175

RESULT 26
 Q8P6F0 PRELIMINARY; PRT; 269 AA.

AC Q8P6F0;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Outer membrane protein.
 GN XAC367.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 CX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RC MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria U.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Secubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL: AE012017; AAM38510.1;
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0005215; F:transporter activity; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR004872; Lipoprotein_9.

DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR InterPro; IPRO00531; TonB_BoxC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
 KM Complete proteome.
 SQ SEQUENCE 269 AA; 29097 MW; 75F2ACE07C28806A CRC64;

Query Match 2.9%; Score 8; DB 16; Length 269;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VARADNMD 244
 DB 230 VARADNMD 237

RESULT 27
 Q87RS3 PRELIMINARY; PRT; 269 AA.
 AC Q87RS3;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Lipoprotein YaeC.
 GN VP0704.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kuokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 RT distinct from that of V. cholerae.";
 RL Lancet 361:743-749(2003).
 DR EMBL; AP05075; BAC58987.1; -
 DR InterPro; IPRO04872; Lipoprotein_9.
 DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Complete proteome.
 SQ SEQUENCE 269 AA; 29070 MW; 5D458BA921862ED8 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 269;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 VAKKYNL 65
 DB 52 VAKKYNL 59

RESULT 28
 Q8CWC8 PRELIMINARY; PRT; 270 AA.
 AC Q8CWC8;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE D-methionine-binding lipoprotein meco precursor.
 GN YAEK OR C0238.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06:HI / CFT073 / ATCC 700928;

FX MEDLINE=22388233; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
 RA Raebio D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AE016755; AAN78730.1; -
 DR InterPro; IPRO04872; Lipoprotein_9.
 DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR InterPro; IPRO04478; YaeC.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR TIGRFAMs; TIGR00363; TIGR00363; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Lipoprotein, Complete proteome.
 SQ SEQUENCE 270 AA; 29318 MW; 2EC4133C671EAE1EC CRC64;

Query Match 2.9%; Score 8; DB 16; Length 270;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 VGNFVYP 120
 DB 106 VGNFVYP 113

RESULT 29
 Q895C6 PRELIMINARY; PRT; 272 AA.
 AC Q895C6;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE ABC transporter substrate binding protein.
 GN CTC01352.
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Massachusetts / B98;
 RX MEDLINE=22457253; PubMed=12552129;
 RA Bruggemann H., Baumer S., Fricke W.F., Wierer A., Liesegang H.,
 RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
 RA Gertschalk G.;
 RT "The genome sequence of Clostridium tetani, the causative agent of
 RT tetanus disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
 DR EMBL; AB015940; AAC5914.1; -
 DR InterPro; IPRO04872; Lipoprotein_9.
 DR InterPro; IPRO00437; Prok_lipoprot_S.
 DR Pfam; PF03180; Lipoprotein_9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Complete proteome.
 SQ SEQUENCE 272 AA; 30259 MW; 36F2AC2CC5368C70 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 272;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 KDSPIYNI 235
 DB 224 KDSPIYNI 231

RESULT 30
 Q7Z3D8 PRELIMINARY; PRT; 277 AA.
 AC Q7Z3D8;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Hypothetical protein DKFZp686L2452 (Fragment).
 GN DKFZp686L2452.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Human endometrium;
 RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
 RA Mewes H.W., Weil B., Amd C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX537968; CAD97932.1; -
 KW Hypothetical protein.
 FT NO_TER
 SQ SEQUENCE 277 AA; 29753 MW; 6A8691B9AE84E7B7 CRC64;

Query Match 2.9%; Score 8; DB 4; Length 277;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 YLEKDSQE 104
 DB 158 YLEKDSQE 165

RESULT 31

ID Q9A1E4 PRELIMINARY; PRT; 281 AA.
 AC Q9A1E4;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Hypothetical protein SPY0319.
 GN SPY0319.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=1196296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezesh S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren O., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RL "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 DR EMBL; AE006497; AA33378.1; -
 DR InterPro; IPR004872; Lipoprotein 9.
 DR Pfam; PF03180; Lipoprotein 9; 1
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 281 AA; 31140 MW; 2D7E7037B0C0C270 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYQDEV 260
 DB 258 KAYQDEV 265

RESULT 32

ID Q8P2K7 PRELIMINARY; PRT; 281 AA.
 AC Q8P2K7;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Hypothetical protein spyM18_0314.
 GN spyM18_0314.
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=186103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108;
 RA Smoot J.C., Barbican K.D., Van Gompel J.U., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.E., Rickielfe S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 RT group A Streptococcus strains associated with acute rheumatic fever
 RT outbreaks."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 DR EMBL; AE009977; AAL97072.1; -
 DR InterPro; IPR004872; Lipoprotein 9.
 DR Pfam; PF03180; Lipoprotein 9; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 281 AA; 31110 MW; 2D907187B0BDC3CB CRC64;

Query Match 2.9%; Score 8; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYQDEV 260
 DB 258 KAYQDEV 265

RESULT 33

ID Q8K8K9 PRELIMINARY; PRT; 281 AA.
 AC Q8K8K9;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Putative lipoprotein (Conserved hypothetical protein).
 GN ATMW OR SPYM3_0233 OR SP31626.
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=198466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22135808; PubMed=12122206;
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schliwert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SSI-1 / Serotype M3;
 RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
 RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
 RA Hayashi H., Hamada S.;
 RT "The genome of invasive Streptococcus pyogenes: a comparative analysis
 RT of S. pyogenes SSI-1, SF370 and MGAS8232."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005146; BAC64721.1; -
 DR InterPro; IPR004872; Lipoprotein 9.

DR InterPro: IPR000437; Prok_Lipoprot_S.
 DR Pfam: PF03180; Lipoprotein_1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Hypothetical protein; Lipoprotein; Complete proteome.
 SQ SEQUENCE 281 AA; 3111 MW; 7D907B87B76D1510 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 KAYOTDEV 260
 DB 258 KAYOTDEV 265

RESULT 34
 Q9P876 PRELIMINARY; PRT; 305 AA.
 AC Q9P876; 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Mitochondrial ADP/ATP carrier isoform 1.
 GN ANC.
 OS Pichia jadinii (Yeast) (Candida utilis).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 CX NCBI_TaxId=4903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 9950;
 RA David C., Trezeguet V., Gonzalez C., Lauguin G.M.;
 RT "Several genes encode the mitochondrial ADP/ATP carrier in the petite
 negative yeast Candida utilis."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 DR EMBL: AJ277099; CAB88027.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0005743; C:mitochondrial inner membrane; IEA.
 DR GO: GO:0005488; F:binding; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR001993; Mitoch carrier.
 DR InterPro: IPR002067; Mit_carrier.
 DR Pfam: PF00153; mito_carri; 3.
 DR PRINTS: PR00926; MITOCARRIER.
 DR PROSITE: PS00215; MITOCH_CARRIER; 2.
 KM Membrane; Transmembrane; Transport.
 SQ SEQUENCE 305 AA; 33202 MW; 9E819EBA0D1AA601 CRC64;

Query Match 2.9%; Score 8; DB 3; Length 305;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAISKTA 36
 DB 21 AAISKTA 28

RESULT 35
 Q9P875 PRELIMINARY; PRT; 305 AA.
 AC Q9P875; 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Mitochondrial ADP/ATP carrier isoform 2.
 GN ANC.
 OS Pichia jadinii (Yeast) (Candida utilis).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 CX NCBI_TaxId=4903;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 9950;
 RA David C., Trezeguet V., Gonzalez C., Lauguin G.M.;
 RT "Several genes encode the mitochondrial ADP/ATP carrier in the petite
 negative yeast Candida utilis."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 DR EMBL: AJ277099; CAB88028.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0005743; C:mitochondrial inner membrane; IEA.
 DR GO: GO:0005488; F:binding; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR001993; Mitoch carrier.
 DR InterPro: IPR002067; Mit_carrier.
 DR Pfam: PF00153; mito_carri; 3.
 DR PRINTS: PR00926; MITOCARRIER.
 DR PROSITE: PS00215; MITOCH_CARRIER; 2.
 KM Membrane; Transmembrane; Transport.
 SQ SEQUENCE 305 AA; 33188 MW; 72F0596160AC7D3C CRC64;

Query Match 2.9%; Score 8; DB 3; Length 305;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAISKTA 36
 DB 21 AAISKTA 28

RESULT 36
 Q8ZSM1 PRELIMINARY; PRT; 347 AA.
 AC Q8ZSM1; 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Hypothetical protein PAE3560.
 GN PAE3560.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 CC Thermoproteales; Pyrobaculum.
 CX NCBI_TaxId=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RC MEDLINE=21664397; PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-U., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 aerophilum."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL: AE009937; AAL65002.1; -
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 347 AA; 38630 MW; 5BCF4154F54420D0 CRC64;

Query Match 2.9%; Score 8; DB 17; Length 347;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 IDVDPLAV 205
 DB 294 IDVDPLAV 301

RESULT 37
 Q8A673 PRELIMINARY; PRT; 351 AA.
 AC Q8A673; 01-JUN-2003 (TREMblrel. 24, Created)
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Putative phosphohydrolases.
 GN BT2013.
 OS Bacteroides thetaiotaomicron.

OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=818;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=1665928;
 RA Xu Y., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
 RA Chang H.C., Hooper L.V., Gordon J.I.,
 RA "A genomic view of the human-bacteroides thetaioamicron symbiosis";
 RL Science 299:2074-2076(2003).
 DR EMBL: AE016934; AAC77120.1;
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR InterPro: IPR004643; M-peptidase.
 DR InterPro: IPR000437; POKR_HIPROTEIN_S.
 DR Pfam: PF00149; Metallophos; 1.
 DR PROSITE: PS00013; PROKR_HIPROTEIN; 1.
 KW Hydrolase; Complete proteome.
 SQ SEQUENCE 351 AA; 39029 MW; 9FB50645491BB86 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 351;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 EKGUNLV 111
 |||||
 Db 142 EKGUNLV 149

RESULT 38
 Q3A18 PRELIMINARY; PRT; 374 AA.
 ID Q3A18;
 AC Q3A18;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE R-phenyllactate dehydratase small subunit.
 GN PLDC.
 OS Clostridium sporogenes.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1509;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 3584;
 RA Dickert S., Pierik A.J., Buckel W.,
 RA "Molecular characterization of (R)-phenyllactate dehydratase and its
 RA activating component A from Clostridium sporogenes";
 RL Submitted (SSP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF420489; AAL18811.1;
 SQ SEQUENCE 374 AA; 43148 MW; 84297814477B60B0 CRC64;

Query Match 2.9%; Score 8; DB 2; Length 374;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 DIVENPKK 183
 |||||
 Db 13 DIVENPKK 20

RESULT 39
 Q9HYE0 PRELIMINARY; PRT; 446 AA.
 ID Q9HYE0;
 AC Q9HYE0;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Probable ATP-dependent RNA helicase.
 GN PA3466.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=29437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mitsuuchi S.D., Watterer P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gader R.J., Goltz L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Lartig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Slier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RA opportunistic pathogen";
 RL Nature 406:959-964(2000).
 CC -1- SIMILARITY: TO DEAD/DEAH BOX HELICASE FAMILY.
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASES.
 DR EMBL: AE004767; AAC6854.1;
 DR PIR: A83213; A83213.
 DR HSPB_Q58083; HVB.
 DR GO: GO:0005622; C:intracellular; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0008026; F:ATP dependent helicase activity; IEA.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR006628; DEAD_box.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR000005; HTHfac.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
 DR PROSITE: PS00041; HTH_ARAC_FAMILY; 1.
 KW ATP-binding; Helicase; RNA-binding; Complete proteome.
 SQ SEQUENCE 446 AA; 48989 MW; 816D6CD046851C6 CRC64;

Query Match 2.9%; Score 8; DB 16; Length 446;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 LARALLL 156
 |||||
 Db 76 LARALLL 83

RESULT 40
 Q99R04 PRELIMINARY; PRT; 512 AA.
 ID Q99R04;
 AC Q99R04;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE NAD-dependent succinate aldehyde dehydrogenases.
 GN ARXK OR BL2368.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiuchi T.,
 RA Sasamoto S., Watanabe A., Idesawa K., Iriyoshi M., Kawasumi K.,
 RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RA "Complete genome sequence of nitrogen-fixing symbiotic bacterium
 RT Bradyrhizobium japonicum USDA110";
 RL DNA Res. 9:189-197(2002).
 DR EMBL: AP005945; BAC48233.1;

DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Complete proteome.
SQ SEQUENCE 512 AA; 55015 MW; 120B682497B83COA CRC64;

Query Match

Best Local Similarity 2.9%; Score 8; DB 16; Length 512;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 135 KDQATIAV 142
Db 367 KDQATIAV 374

Search completed: June 16, 2004, 11:17:15
Job time : 69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:08:47 / Search time 61 seconds
(without alignments)
1278.412 Million cell updates/sec

Title: US-10-018-672-2

Perfect score: 276
Sequence: 1 MNRGKINGICALASGIALAG.....TDEVEAEAKKQKPDGVYKGM 276

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 1586107 segs, 282547505 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database: A_Geneseq_29Jan04:*

1: Geneseq1980s:*\n2: Geneseq1990s:*\n3: Geneseq2000s:*\n4: Geneseq2001s:*\n5: Geneseq2002s:*\n6: Geneseq2003as:*\n7: Geneseq2003bs:*\n8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	276	100.0	276	4	AAB20104
2	243	88.0	276	5	AAOI7582
3	243	88.0	276	6	ABU35462
4	19	6.9	272	5	AAOI7813
5	19	6.9	273	4	AAU35474
6	19	6.9	273	5	ABR76073
7	19	6.9	273	5	ABR76073
8	19	6.9	273	5	ABR76073
9	16	5.8	270	5	ABR76095
10	16	5.8	271	6	ABU40467
11	14	5.1	20	5	ABR76095
12	14	4.7	270	6	ABU24914
13	14	4.3	276	6	ABU39463
14	14	4.0	67	4	AAU70412
15	14	4.0	271	6	ABU27902
16	11	4.0	271	6	ABU48271
17	10	3.6	276	6	ABU16868
18	10	3.6	276	6	ABU16868
19	9	3.3	125	6	ABR53515
20	9	3.3	241	6	ABU16864
21	9	3.3	256	6	ABU33203
22	9	3.3	262	6	ABU23447
23	9	3.3	270	6	ABU27880
24	9	3.3	272	6	ABM67576
25	9	3.3	273	5	ABR48120

26	9	3.3	273	6	ABU32634	ABU32634 Protein e
27	9	3.3	275	6	ABU49250	ABU49250 Protein e
28	9	3.3	291	7	ADG97607	ADG97607 E. faeciu
29	9	3.3	296	6	ADA35637	ADA35637 Acinetoba
30	8	2.9	96	4	AAU18017	AAU18017 Human imm
31	8	2.9	96	4	ABR10497	ABR10497 Human CDN
32	8	2.9	96	5	ABP67084	ABP67084 Human pol
33	8	2.9	96	7	AD331641	AD331641 Human nov
34	8	2.9	152	4	AAO05643	AAO05643 Human pol
35	8	2.9	240	4	AAU36063	AAU36063 Klebsiell
36	8	2.9	263	5	ABP26478	ABP26478 Streptoco
37	8	2.9	263	6	ABU25052	ABU25052 Protein e
38	8	2.9	265	6	ABU23277	ABU23277 Protein e
39	8	2.9	269	7	ADC96816	ADC96816 E. faeciu
40	8	2.9	270	6	AAU21415	AAU21415 Protein e
41	8	2.9	271	4	AAU34457	AAU34457 E. coli c
42	8	2.9	271	6	ABU31676	ABU31676 Protein e
43	8	2.9	271	6	ABU15367	ABU15367 Protein e
44	8	2.9	281	6	ABU46465	ABU46465 Protein e
45	8	2.9	282	6	ABU29987	ABU29987 Protein e
46	8	2.9	295	6	ABU19780	ABU19780 Protein e
47	8	2.9	379	4	ABR10313	ABR10313 Human CDN
48	8	2.9	379	5	ABP66900	ABP66900 Human pol
49	8	2.9	739	5	AAU68310	AAU68310 Human sem
50	8	2.9	744	5	AAU68310	AAU68310 Human sem
51	8	2.9	758	4	ABR58337	ABR58337 Human sem
52	8	2.9	766	5	AAU68312	AAU68312 Human sem
53	8	2.9	771	5	AAU68311	AAU68311 Human sem
54	8	2.9	838	5	AAU68302	AAU68302 Human sem
55	8	2.9	843	5	AAU68301	AAU68301 Human sem
56	8	2.9	865	5	AAU68304	AAU68304 Human sem
57	8	2.9	870	5	AAU68303	AAU68303 Human sem
58	8	2.9	9	5	ABR76087	ABR76087 Haemophil
59	8	2.5	9	5	ABR76093	ABR76093 Haemophil
60	8	2.5	9	5	ABR76086	ABR76086 Haemophil
61	8	2.5	9	5	ABR76092	ABR76092 Haemophil
62	8	2.5	73	2	AAU20141	AAU20141 H. pylori
63	8	2.5	78	2	AAU38625	AAU38625 Streptoco
64	8	2.5	92	4	AAU25520	AAU25520 Human pro
65	8	2.5	98	5	ABP07971	ABP07971 Human ORF
66	8	2.5	105	4	AAU81880	AAU81880 S. epider
67	8	2.5	114	7	ADCO0063	ADCO0063 Enteroha
68	8	2.5	114	7	ADP99989	ADP99989 Enteroha
69	8	2.5	114	7	ADCO0470	ADCO0470 Enteroha
70	8	2.5	114	7	ADCO0739	ADCO0739 Enteroha
71	8	2.5	114	7	ADCO0848	ADCO0848 Enteroha
72	8	2.5	114	7	ADCO1030	ADCO1030 Enteroha
73	8	2.5	114	7	ADCO0319	ADCO0319 Enteroha
74	8	2.5	116	5	ABP38486	ABP38486 Stephyloc
75	8	2.5	117	5	ABP30711	ABP30711 Streptoco
76	8	2.5	131	3	AAU57230	AAU57230 Human pro
77	8	2.5	132	3	AAU44014	AAU44014 Human can
78	8	2.5	151	5	ABP30811	ABP30811 Streptoco
79	8	2.5	153	2	AAU20166	AAU20166 Helicobac
80	8	2.5	159	4	AAU03580	AAU03580 Human pro
81	8	2.5	161	2	AAU09625	AAU09625 P. aerugi
82	8	2.5	172	5	ABP28372	ABP28372 Streptoco
83	8	2.5	173	4	AAU29665	AAU29665 Novel hum
84	8	2.5	174	7	ADCO6159	ADCO6159 Human GPC
85	8	2.5	178	7	ADCO6153	ADCO6153 Human GPC
86	8	2.5	186	4	ABR69789	ABR69789 Human bra
87	8	2.5	194	4	AAU14092	AAU14092 Peptide #
88	8	2.5	194	4	ABR33037	ABR33037 Peptide #
89	8	2.5	194	4	AAU26499	AAU26499 Peptide #
90	8	2.5	194	4	ABR27866	ABR27866 Human pep
91	8	2.5	194	4	ABR18510	ABR18510 Protein #
92	8	2.5	194	4	AAU66222	AAU66222 Human bon
93	8	2.5	194	4	AAU38336	AAU38336 Human bra
94	8	2.5	194	4	AAU47890	AAU47890 Human liv
95	8	2.5	194	4	AAU01830	AAU01830 Peptide #
96	8	2.5	194	5	ABG35872	ABG35872 Human pep
97	8	2.5	197	4	ABG16733	ABG16733 Novel hum
98	8	2.5	199	2	AAU60986	AAU60986 A 22-26 x

99	2.5	214	4	ABG00371	Abg00371	Novel hum	172	350	1	ABP2053	ABP2053	Outer mem
100	2.5	220	6	ABU41510	Abu41510	Protein e	173	350	4	ABP5179	ABP5179	Protein a
101	2.5	220	7	ADC39190	Adc39190	Novel hum	174	352	4	AAU34460	AAU34460	E. coli c
102	2.5	222	4	ABG27431	Abg27431	Novel hum	175	352	6	ABM67622	ABM67622	Phototrab
103	2.5	222	5	ABP74106	Abp74106	Human TRI	176	355	7	ADC86179	ADC86179	Human GPC
104	2.5	226	2	ABP96827	Abp96827	P. aerugi	177	356	4	AAE03581	AAE03581	Human pro
105	2.5	228	5	ABP43205	Abp43205	Human ova	178	411	7	ADC46587	ADC46587	Chlamydom
106	2.5	241	6	ABP07136	Abp07136	Alliolooc	179	419	2	ABP25465	ABP25465	Mycobacte
107	2.5	244	4	ABE80307	AbE80307	Corynbac	180	419	6	ABG81240	ABG81240	Mycobacte
108	2.5	244	5	ABG80339	Abg80339	C. glutam	181	419	6	ABU37016	ABU37016	Protein e
109	2.5	247	6	ABG26453	Abg26453	Protein e	182	419	6	ABU34873	ABU34873	Protein e
110	2.5	256	6	ABU39703	Abu39703	Protein e	183	424	4	AAAG90552	AAAG90552	C. glutam
111	2.5	260	4	AAU36517	AAu36517	Pseudomon	184	424	7	ADD13573	ADD13573	C. glutam
112	2.5	260	6	ABU38920	Abu38920	Protein e	185	429	6	ABR41715	ABR41715	Human DIP
113	2.5	261	6	ABU43388	Abu43388	Protein e	186	438	4	AAE03582	AAE03582	Human pro
114	2.5	262	6	ABU49395	Abu49395	Protein e	187	460	4	AAAG90469	AAAG90469	C. glutam
115	2.5	262	6	ABU19705	Abu19705	Protein e	188	468	4	AAE03574	AAE03574	Human pro
116	2.5	263	4	ABG27056	Abg27056	Novel hum	189	475	2	AAK31046	AAK31046	Kat DIB d
117	2.5	265	6	ABU39208	Abu39208	Protein e	190	477	2	AAK21082	AAK21082	Dopamine
118	2.5	268	2	AAK90682	AAK90682	Rat calde	191	477	2	AAK79381	AAK79381	Dopamine
119	2.5	268	6	ABU48077	Abu48077	Protein e	192	477	2	AAW40802	AAW40802	Human DS
120	2.5	268	6	ABU27644	Abu27644	Protein e	193	477	2	ABM56350	ABM56350	Non-endog
121	2.5	268	6	ABU21332	Abu21332	Protein e	194	477	6	ABP81808	ABP81808	Human dop
122	2.5	268	6	ABU47067	Abu47067	Protein e	195	479	2	AAK22546	AAK22546	Truncated
123	2.5	268	6	ABU14874	Abu14874	Protein e	196	480	5	ABG78365	ABG78365	Human pro
124	2.5	268	6	ABU45850	Abu45850	Protein e	197	496	6	ABG81968	ABG81968	S. epider
125	2.5	269	6	ABU30644	Abu30644	Protein e	198	496	6	ABU45063	ABU45063	Protein e
126	2.5	269	6	ABU41046	Abu41046	Protein e	199	505	4	ABG15313	ABG15313	Novel hum
127	2.5	269	6	ABU31515	Abu31515	Protein e	200	506	5	ABP39538	ABP39538	Staphyloc
128	2.5	270	6	ABM69136	ABm69136	Phototrab	201	507	4	AAE03576	AAE03576	Human pro
129	2.5	270	6	ABU17607	Abu17607	Protein e	202	519	5	ABP65972	ABP65972	Human pol
130	2.5	271	2	AAU10962	AAU10962	H. pylori	203	527	4	ABM61819	ABM61819	Drosophila
131	2.5	271	2	AAU10998	AAU10998	H. pylori	204	556	5	ABP73756	ABP73756	Candida a
132	2.5	271	2	AAK89937	AAK89937	Anti-Igen 1	205	575	5	AAO17811	AAO17811	H. influenza
133	2.5	271	2	AAW89833	AAW89833	Protein e	206	589	4	AAE03577	AAE03577	Human pro
134	2.5	271	2	AAW90023	AAW90023	Expressed	207	611	4	AAK96570	AAK96570	Putative
135	2.5	271	4	AAU35867	AAU35867	Helicobac	208	632	6	ABU26147	ABU26147	Protein e
136	2.5	271	4	AAU36031	AAU36031	Helicobac	209	652	2	AAW80610	AAW80610	S. pneumo
137	2.5	271	6	ABU50250	Abu50250	Protein e	210	652	4	AAU38080	AAU38080	Streptococ
138	2.5	271	6	ABU50643	Abu50643	Protein e	211	652	4	AAU37724	AAU37724	Streptococ
139	2.5	271	6	ABU43419	Abu43419	Protein e	212	652	6	ABU01536	ABU01536	S. pneumo
140	2.5	271	6	ABU31110	Abu31110	Protein e	213	652	6	ABU46049	ABU46049	Protein e
141	2.5	272	6	ABU24666	Abu24666	Protein e	214	668	4	ABM61757	ABM61757	Drosophila
142	2.5	274	6	ABU50032	Abu50032	Protein e	215	670	4	ABM71706	ABM71706	Drosophila
143	2.5	276	5	ABM49303	ABM49303	Listeria	216	675	6	ABU19181	ABU19181	Protein e
144	2.5	276	6	ABU32999	Abu32999	Protein e	217	689	4	ABM59812	ABM59812	Drosophila
145	2.5	280	2	AAW20909	AAW20909	H. pylori	218	689	4	ABM68912	ABM68912	Drosophila
146	2.5	286	2	AAU85996	AAU85996	S. pneumo	219	690	7	ABM00766	ABM00766	Helicobac
147	2.5	286	5	ABM53624	ABM53624	Lactococc	220	691	2	AAW73027	AAW73027	Helicobac
148	2.5	286	5	ABM53625	ABM53625	Lactococc	221	691	2	AAW73032	AAW73032	Helicobac
149	2.5	293	3	ADC39192	Adc39192	Novel hum	222	691	7	ABM00781	ABM00781	Helicobac
150	2.5	298	3	AAU71032	AAU71032	Human ade	223	692	4	AAK93225	AAK93225	C. glutam
151	2.5	298	3	AAU71033	AAU71033	Human ade	224	717	1	AAK91933	AAK91933	BI antigen
152	2.5	298	4	AAU01199	AAU01199	Human ade	225	757	4	AAE03583	AAE03583	Human pro
153	2.5	298	4	AAU01200	AAU01200	Human ade	226	766	4	AAK93402	AAK93402	Human pol
154	2.5	298	4	AAU39641	AAU39641	Human pol	227	852	6	ABP79237	ABP79237	N. gonorr
155	2.5	298	5	AAO18516	AAO18516	Human ins	228	852	6	ABU21216	ABU21216	Human pro
156	2.5	298	5	AAU10379	AAU10379	Human ade	229	852	6	ABU37147	ABU37147	Protein e
157	2.5	298	5	AAU10380	AAU10380	Human ade	230	852	6	ABU37977	ABU37977	Protein e
158	2.5	302	5	ABP38326	ABP38326	Staphyloc	231	858	4	AAK27282	AAK27282	Human ADA
159	2.5	307	5	ABM53940	ABM53940	Lactococc	232	860	4	AAK92710	AAK92710	C. glutam
160	2.5	309	3	AAU18898	AAU18898	A. maize c	233	871	3	AAU77289	AAU77289	Streptococ
161	2.5	310	4	AAE03575	AAE03575	Human pro	234	871	3	AAU78841	AAU78841	KSG-ATG 1
162	2.5	311	6	ABU05082	Abu05082	N. mening	235	875	5	ABM48886	ABM48886	Listeria
163	2.5	311	6	ABU37851	Abu37851	Protein e	236	875	5	ABU32754	ABU32754	Protein e
164	2.5	316	6	ABP77807	ABP77807	N. gonorr	237	908	4	AAE03572	AAE03572	Human pro
165	2.5	317	6	AAE03579	AAE03579	Human pro	238	959	4	ABM69647	ABM69647	Human met
166	2.5	317	6	AAE03579	AAE03579	Human pro	239	1036	6	ABU49046	ABU49046	Protein e
167	2.5	323	4	AAU41427	AAU41427	Human pol	240	1120	5	ABG76502	ABG76502	DNA encod
168	2.5	325	4	ABG15423	ABG15423	Novel hum	241	1128	6	ABU41214	ABU41214	Protein e
169	2.5	336	4	ABG21669	ABG21669	Novel hum	242	1172	4	ABM66466	ABM66466	Drosophila
170	2.5	337	4	ABP92804	ABP92804	Microsom	243	1175	4	AAU55065	AAU55065	Protein e
171	2.5	337	6	ABP92833	ABP92833	Othosomy	244	1175	5	ABP73177	ABP73177	Candida a

245	7	2.5	1178	6	ABU24510	Protein e	318	6	2.2	34	2	AAR40887	Aar40887 Sequence
246	7	2.5	1233	4	AAG30558	Agg30558 C. glutam	319	6	2.2	34	2	AAR40888	Aar40888 Sequence
247	7	2.5	1233	7	ADD13972	Add13972 C. glutam	320	6	2.2	34	2	AAR40886	Aar40886 Sequence
248	7	2.5	1239	5	AAO21545	Aao21545 Human pol	321	6	2.2	34	2	AAR35238	Aar35238 Human leu
249	7	2.5	1241	3	AAB44262	Aab44262 Human ORF	322	6	2.2	34	2	AAM35240	Aam35240 Human leu
250	7	2.5	1263	2	AAW22052	Aaw22052 DNA polym	323	6	2.2	34	2	AAM35241	Aam35241 Human leu
251	7	2.5	1482	3	AAAY84727	Aay84727 Antino aci	324	6	2.2	34	2	AAW35237	Aaw35237 Human leu
252	7	2.5	1488	3	AAAY84725	Aay84725 Antino aci	325	6	2.2	34	2	AAW35239	Aaw35239 Human leu
253	7	2.5	1509	3	AAAY84728	Aay84728 Antino aci	326	6	2.2	35	6	ABP68391	Abp68391 Human col
254	7	2.5	1517	3	AAAY84726	Aay84726 Antino aci	327	6	2.2	37	4	ABG28355	Abg28355 Novel hum
255	7	2.5	1615	4	AAU03536	Aau03536 Human pro	328	6	2.2	40	3	ABR12869	Abri2869 Degenerat
256	7	2.5	1972	2	AAW81171	Aaw81171 Human BAZ	329	6	2.2	42	2	AAW41874	Aaw41874 Peptide u
257	7	2.5	1972	6	ABR64241	Abre64241 Angiogene	330	6	2.2	42	5	ABB05820	Abb05820 Amb a I M
258	7	2.5	2154	6	ABU38771	Abu38771 Protein e	331	6	2.2	42	5	ABB05892	Abb05892 Amb a I M
259	7	2.5	3241	6	ABU22000	Abu22000 Protein e	332	6	2.2	43	2	AAR54448	Aar54448 Amb a I.1
260	7	2.5	4688	6	ABU48941	Abu48941 Protein e	333	6	2.2	43	2	AAW02503	Aaw02503 Peptide c
261	7	2.5	6310	5	ABU39869	Abu39869 Protein e	334	6	2.2	43	4	AAW83618	Aaw83618 Human imm
262	6	2.2	7	5	ABBT76084	Abbt76084 Haemophil	335	6	2.2	43	6	ABR83844	Abri83844 Drosophil
263	6	2.2	8	4	AAU04099-	Aau04099 Human CLA	336	6	2.2	43	6	ABR83843	Abri83843 Drosophil
264	6	2.2	9	5	ABB76097	Abbt76097 Haemophil	337	6	2.2	44	2	AAR54444	Aar54444 Amb a I.1
265	6	2.2	9	5	ABB76091	Abbt76091 Haemophil	338	6	2.2	44	2	AAW02499	Aaw02499 Peptide c
266	6	2.2	10	5	ABG98527	Abg98527 F. protein	339	6	2.2	45	2	AAW02499	Aaw02499 Peptide c
267	6	2.2	10	5	ABG98528	Abg98528 F. protein	340	6	2.2	45	2	AAW41873	Aaw41873 Human sec
268	6	2.2	10	5	ABG98531	Abg98531 F. protein	341	6	2.2	45	3	ABR25733	Abri25733 Human sec
269	6	2.2	10	5	ABG98530	Abg98530 F. protein	342	6	2.2	45	3	ABR05819	Abri05819 Amb a TA
270	6	2.2	10	5	ABG98529	Abg98529 F. protein	343	6	2.2	45	5	ABB05891	Abb05891 Amb a TA
271	6	2.2	11	4	AAW21957	Aaw21957 Glycopept	344	6	2.2	48	3	AAG04829	Aag04829 Arabidops
272	6	2.2	11	4	AAU04100	Aau04100 Human CLA	345	6	2.2	49	4	ABG24210	Abg24210 Novel hum
273	6	2.2	11	5	ABG67601	Abg67601 Human ADP	346	6	2.2	50	4	AAU56678	Aau56678 Human ORF
274	6	2.2	11	6	ADA23719	Ada23719 Alzheimer	347	6	2.2	50	5	ABP02306	Abp02306 Human ORF
275	6	2.2	12	5	ABT76098	Abbt76098 Haemophil	348	6	2.2	50	6	ABE53197	Abri53197 Propionib
276	6	2.2	12	5	ABT76099	Abbt76099 Haemophil	349	6	2.2	50	7	AOC01216	Aoc01216 Enterococ
277	6	2.2	13	4	ADD23878	Add23878 Breast ca	350	6	2.2	52	4	AAW91855	Aaw91855 Human dig
278	6	2.2	13	4	AAW93371	Aaw93371 Vaccine r	351	6	2.2	52	4	AAU22484	Aau22484 Novel hum
279	6	2.2	14	4	AAW78359	Aaw78359 NAM mltog	352	6	2.2	52	7	ADB32324	Adb32324 Human nov
280	6	2.2	14	4	AAW72007	Aaw72007 Internal	353	6	2.2	54	2	AAW77560	Aaw77560 Staphyloc
281	6	2.2	15	5	AAR37897	Aar37897 Formylpep	354	6	2.2	54	2	AAW78358	Aaw78358 M. arthri
282	6	2.2	16	5	ABBT7564	Abbt7564 Linker IT	355	6	2.2	54	4	AAU63866	Aau63866 Propionib
283	6	2.2	17	1	AAPE83373	Aape83373 Immunosup	356	6	2.2	54	6	ABE59985	Abri59985 Propionib
284	6	2.2	17	1	AAPE83370	Aape83370 Immunosup	357	6	2.2	57	2	AAAY5826	Aay5826 S. pneumo
285	6	2.2	17	1	AAPE83372	Aape83372 Immunosup	358	6	2.2	57	6	AAE14831	Aae14831 Human int
286	6	2.2	17	2	AAAT07352	Aaat07352 RAE26.15	359	6	2.2	57	6	AAE14833	Aae14833 Human int
287	6	2.2	17	2	AAW41907	Aaw41907 Peptide u	360	6	2.2	59	4	AAU46397	Aau46397 Propionib
288	6	2.2	17	2	AAW58735	Aaw58735 Allergen	361	6	2.2	59	4	AAU52789	Aau52789 Propionib
289	6	2.2	17	5	ABB05885	Abb05885 Ragweed p	362	6	2.2	59	6	ABY42816	Abi42816 Propionib
290	6	2.2	20	2	AAAB05927	Aaab05927 Ragweed p	363	6	2.2	59	6	ABY49308	Abi49308 Propionib
291	6	2.2	20	2	AAAB83334	Aaab83334 S. aureus	364	6	2.2	61	2	AAW38547	Aaw38547 Streptoco
292	6	2.2	20	2	AAW85291	Aaw85291 Helper T-	365	6	2.2	61	4	AAW89456	Aaw89456 Human imm
293	6	2.2	21	1	AAAP20042	Aap20042 Haptlen re	366	6	2.2	63	4	ABE66776	Abi66776 Drosophil
294	6	2.2	21	1	AAAP30017	Aap30017 Human lym	367	6	2.2	64	4	AAU45989	Aau45989 Propionib
295	6	2.2	21	2	AAAR54428	Aaar54428 AMB 3-4.1	368	6	2.2	64	6	AAW42508	Aaw42508 Propionib
296	6	2.2	22	2	AAAR54377	Aaar54377 RAE 66.1-	369	6	2.2	66	7	AAI02898	Aai02898 Fragment
297	6	2.2	22	2	AAW02461	Aaw02461 RAE 66.1-	370	6	2.2	66	7	ADA07833	Ada07833 Human sec
298	6	2.2	24	2	AAAR54432	Aaar54432 AMB 3-4.1	371	6	2.2	67	5	AAU98428	Aau98428 Cadium-re
299	6	2.2	24	2	AAW02411	Aaw02411 AMB 3-4.1	372	6	2.2	69	3	AAW39334	Aaw39334 Arabidops
300	6	2.2	24	2	AAW02494	Aaw02494 AMB 3-4.1	373	6	2.2	71	3	AAW61078	Aaw61078 Arabidops
301	6	2.2	25	2	AAAR54376	Aaar54376 RAE 66.1-	374	6	2.2	72	2	AAW08064	Aaw08064 Gas vesic
302	6	2.2	25	2	AAW02460	Aaw02460 RAE 66.1-	375	6	2.2	73	5	ABP09239	Abp09239 Human ORF
303	6	2.2	27	2	AAAR54427	Aaar54427 AMB 3-5.1	376	6	2.2	75	4	AAW84743	Aaw84743 Human imm
304	6	2.2	27	2	AAW02410	Aaw02410 AMB 3-5.1	377	6	2.2	75	4	AAO06742	Aao06742 Human pol
305	6	2.2	27	2	AAW69259	Aaw69259 DNS Prote	378	6	2.2	76	4	ABG02551	Abg02551 Novel hum
306	6	2.2	28	2	AAAR54375	Aaar54375 RAE 66.1-	379	6	2.2	77	3	AAW11804	Aaw11804 Arabidops
307	6	2.2	28	2	AAW02459	Aaw02459 RAE 66.1	380	6	2.2	78	4	AAU22981	Aau22981 Novel hum
308	6	2.2	28	7	ADFA8198	Adfa8198 Putative	381	6	2.2	78	5	ABP03410	Abp03410 Human ORF
309	6	2.2	29	4	ABG24695	Abg24695 Novel hum	382	6	2.2	79	2	AAW75397	Aaw75397 Human bra
310	6	2.2	29	6	ADA04800	Ada04800 IGF-IR b1	383	6	2.2	79	4	AAU39289	Aau39289 Propionib
311	6	2.2	30	2	AAW41878	Aaw41878 Peptide u	384	6	2.2	79	6	ABR35808	Abri35808 Propionib
312	6	2.2	30	5	ABP29573	Abp29573 Streptoco	385	6	2.2	80	6	ADB07976	Adi07976 Alloloco
313	6	2.2	30	5	ABW05824	Abw05824 AGE MAJOR	386	6	2.2	81	4	AAO03486	Aao03486 Human pol
314	6	2.2	30	5	ABW05896	Abw05896 AGE MAJOR	387	6	2.2	81	6	ABU24670	Abu24670 Protein e
315	6	2.2	34	2	AAR40885	Aar40885 Sequence	388	6	2.2	82	4	AAU52720	Aau52720 Propionib
316	6	2.2	34	2	AAR40881	Aar40881 N-termina	389	6	2.2	82	5	ABP09448	Abp09448 Human ORF
317	6	2.2	34	2	AAR40884	Aar40884 Sequence	390	6	2.2	82	6	ABW49239	Abw49239 Propionib

391	6	2.2	82	6	ABU07664	Human hap	464	6	2.2	115	6	ABU01743	S. pneumo
392	6	2.2	82	7	ADCC00634	Enterococ	465	6	2.2	115	6	ABR83559	Tagged To
393	6	2.2	84	4	AAU46356	Proteobact	466	6	2.2	116	6	ABU45107	Protein e
394	6	2.2	84	4	AAU47491	Proteobact	467	6	2.2	116	7	ADCC89438	Ribosomal
395	6	2.2	84	5	ABP09893	Human ORF	468	6	2.2	117	2	AAK52053	Heavy cha
396	6	2.2	84	5	ABR49563	Listeria	469	6	2.2	117	2	ABP35225	Human ORF
397	6	2.2	84	5	ABM42875	Proteobact	470	6	2.2	117	7	ADCC88479	Ribosomal
398	6	2.2	84	6	ABM44010	Proteobact	471	6	2.2	118	2	AAU08460	Rat Ptx2
399	6	2.2	84	6	ABU32482	Protein e	472	6	2.2	118	7	ADCC7894	Ribosomal
400	6	2.2	85	4	AAU21455	Human nov	473	6	2.2	119	1	AAU20008	Hybrid hu
401	6	2.2	87	3	AAU04827	Arabisdops	474	6	2.2	119	4	AAU67797	Proteobact
402	6	2.2	87	3	AAU16177	Human nov	475	6	2.2	119	5	AAE19376	Human rep
403	6	2.2	87	6	ABU55246	Human nov	476	6	2.2	119	5	ABE78137	Consensus
404	6	2.2	88	5	ABP42513	Human ova	477	6	2.2	119	5	ABR05755	Reprolysi
405	6	2.2	88	5	ABU55661	Human nov	478	6	2.2	119	5	ABM64316	Proteobact
406	6	2.2	89	3	ABG57673	Arabisdops	479	6	2.2	119	6	ABK36698	HMM deriv
407	6	2.2	89	4	AAU16592	Human nov	480	6	2.2	119	6	ABU08384	Consensus
408	6	2.2	89	4	AAU44672	Proteobact	481	6	2.2	120	2	AAW08366	RSV F pro
409	6	2.2	89	4	ABM16429	Human nov	482	6	2.2	120	2	ABG12411	Novel hum
410	6	2.2	89	6	ABM41191	Human nov	483	6	2.2	120	5	ABE04303	Recombin
411	6	2.2	89	6	ABU55661	Human ova	484	6	2.2	121	4	AAU00446	Human ORF
412	6	2.2	93	3	AAU1070	Arabisdops	485	6	2.2	121	5	ABP09591	Human pol
413	6	2.2	93	3	AAU1070	Arabisdops	486	6	2.2	121	5	ABE90187	Human pol
414	6	2.2	93	6	ABR83556	Escherich	487	6	2.2	122	4	AAU42386	Proteobact
415	6	2.2	93	7	ADCC3781	Human nov	488	6	2.2	122	6	AAU38905	Proteobact
416	6	2.2	94	5	ABG69875	Human sec	489	6	2.2	122	7	ADCC88432	Ribosomal
417	6	2.2	94	6	ABP75798	Human sec	490	6	2.2	123	3	AAU48509	Arabisdops
418	6	2.2	95	3	AAU39605	Arabisdops	491	6	2.2	123	3	AAU41908	Arabisdops
419	6	2.2	95	3	AAU1069	Arabisdops	492	6	2.2	123	4	AAU17154	Peptide #
420	6	2.2	95	6	ABP77850	N. gonorr	493	6	2.2	123	4	ABK36155	Peptide #
421	6	2.2	97	4	AAU86838	Human imm	494	6	2.2	123	4	AAU29646	Peptide #
422	6	2.2	98	4	AAU66987	Proteobact	495	6	2.2	123	4	ABK30964	Peptide #
423	6	2.2	98	4	ABG06348	Novel hum	496	6	2.2	123	4	ABK21542	Protein #
424	6	2.2	98	6	ABM63506	Proteobact	497	6	2.2	123	4	AAU69324	Human bon
425	6	2.2	99	4	AAU40436	Proteobact	498	6	2.2	123	4	AAU70858	Human bon
426	6	2.2	99	6	ABP31616	Human ORF	499	6	2.2	123	4	AAU56940	Human bon
427	6	2.2	99	6	ABM36955	Proteobact	500	6	2.2	123	4	ABG51000	Human liv
428	6	2.2	100	6	AAU66991	Proteobact	501	6	2.2	123	4	AAU04849	Peptide #
429	6	2.2	100	6	ABM63510	Proteobact	502	6	2.2	123	5	ABG38941	Human pep
430	6	2.2	100	7	ADCC8456	Protein e	503	6	2.2	125	3	AAU27567	Arabisdops
431	6	2.2	100	7	ADCC8456	Ribosomal	504	6	2.2	125	3	AAU13561	Arabisdops
432	6	2.2	101	2	AAU74194	Human pro	505	6	2.2	125	3	AAU05365	Arabisdops
433	6	2.2	101	4	ABG25250	Arabisdops	506	6	2.2	125	4	AAU67588	Proteobact
434	6	2.2	101	4	ABG26432	Novel hum	507	6	2.2	125	6	ABE64107	Proteobact
435	6	2.2	101	5	ABR06357	Human zin	508	6	2.2	126	3	AAU52850	Arabisdops
436	6	2.2	102	2	AAU04569	Human prod	509	6	2.2	126	3	AAU52840	Arabisdops
437	6	2.2	102	2	AAU73871	Arabisdops	510	6	2.2	127	5	AAU48508	Arabisdops
438	6	2.2	102	2	AAU1068	Arabisdops	511	6	2.2	128	6	ABU23794	Protein e
439	6	2.2	102	3	AAU39604	Arabisdops	512	6	2.2	128	6	ABU24054	Protein e
440	6	2.2	102	5	ABP27936	Streptococ	513	6	2.2	130	3	AAU25249	Secreted
441	6	2.2	103	5	ABG93246	C. albica	514	6	2.2	131	2	AAU59719	Secreted
442	6	2.2	103	5	ABG93328	C. albica	515	6	2.2	131	4	ABG02740	Novel hum
443	6	2.2	103	5	ABP73948	Candida a	516	6	2.2	132	4	ABG27850	Novel hum
444	6	2.2	104	4	AAU50126	Proteobact	517	6	2.2	132	5	ABP32084	Human ORF
445	6	2.2	104	4	ABU51226	Human ner	518	6	2.2	132	6	ABU52320	Human GPC
446	6	2.2	104	4	ABU49860	Listeria	519	6	2.2	133	2	AAU07877	Human sec
447	6	2.2	104	4	ABU46645	Proteobact	520	6	2.2	133	4	ABG03888	Novel hum
448	6	2.2	107	6	ADU43275	Actinocoba	521	6	2.2	133	2	AAU38829	Neisseria
449	6	2.2	108	3	AAU39333	Arabisdops	522	6	2.2	135	2	AAU38930	Neisseria
450	6	2.2	109	1	AAU61525	Sequence	523	6	2.2	135	2	AAU38927	Neisseria
451	6	2.2	109	1	AAU61525	Sequence	524	6	2.2	135	2	AAU38927	Neisseria
452	6	2.2	110	4	ABU48332	Drosophil	525	6	2.2	135	3	AAU75017	Neisseria
453	6	2.2	111	3	AAU36772	Human sec	526	6	2.2	135	3	AAU75020	Neisseria
454	6	2.2	111	3	AAU36772	Human sec	527	6	2.2	135	3	AAU57671	Arabisdops
455	6	2.2	111	4	AAU47868	Proteobact	528	6	2.2	135	4	AAU03521	Human gen
456	6	2.2	111	5	AAU47868	Proteobact	529	6	2.2	135	4	AAU23109	Novel hum
457	6	2.2	111	5	AAU47868	Proteobact	530	6	2.2	135	4	AAU010323	Human pol
458	6	2.2	111	6	ABM43387	Proteobact	531	6	2.2	135	4	AAU44634	Proteobact
459	6	2.2	111	7	ABM74341	DNA clone	532	6	2.2	135	5	ABG63423	Human alb
460	6	2.2	114	3	AAU42766	Human ORF	533	6	2.2	135	6	ABP80909	N. gonorr
461	6	2.2	114	4	AAU28400	Human G P	534	6	2.2	135	6	ABM41153	Proteobact
462	6	2.2	114	5	ABG60688	Novel G P	535	6	2.2	135	6	ABM71769	Stephyloc
463	6	2.2	115	4	ABU70709	Drosophil	536	6	2.2	136	3	AAU78979	Silkworm

537	6	2.2	137	4	AAO12794	AAO12794 Human pol	610	6	2.2	154	5	ABG79185	ABg79185 Human rib
538	6	2.2	137	5	ABP08419	ABP08419 Human CRF	611	6	2.2	155	2	AAy29570	AAy29570 Human lun
539	6	2.2	137	5	AAE25600	AAE25600 Human ino	612	6	2.2	155	3	AAAB44116	AAAB44116 Human can
540	6	2.2	138	4	AAO01261	AAO01261 Human pol	613	6	2.2	155	3	AAAB44451	AAAB44451 Human lun
541	6	2.2	138	4	AAU53569	AAU53569 Propionib	614	6	2.2	155	4	AAU65508	AAU65508 Propionib
542	6	2.2	138	6	ABM50088	ABM50088 Propionib	615	6	2.2	155	4	AAE13792	AAE13792 Human lun
543	6	2.2	139	3	AAO10263	AAO10263 Arabinidops	616	6	2.2	155	4	AAAG82274	AAAG82274 S. epider
544	6	2.2	139	3	AAO05364	AAO05364 Arabinidops	617	6	2.2	155	6	ABM62027	ABM62027 Propionib
545	6	2.2	139	3	AAO13560	AAO13560 Arabinidops	618	6	2.2	155	7	ADP65422	ADP65422 Human lun
546	6	2.2	139	3	AAO29335	AAO29335 Arabinidops	619	6	2.2	155	7	ADP87676	ADP87676 Human lun
547	6	2.2	139	3	AAO27566	AAO27566 Arabinidops	620	6	2.2	156	2	AAAB6743	AAAB6743 Biotin ca
548	6	2.2	139	4	AAAB76796	AAAB76796 Coryneb	621	6	2.2	156	4	AAU36469	AAU36469 Pseudomon
549	6	2.2	139	4	AAAB80210	AAAB80210 Coryneb	622	6	2.2	156	4	AAU94742	AAU94742 Human rep
550	6	2.2	139	5	ABG80342	ABG80342 C. glutam	623	6	2.2	156	6	ABG22975	ABG22975 Novel hum
551	6	2.2	140	3	AAO52949	AAO52949 Arabinidops	624	6	2.2	156	6	ABU38815	ABU38815 Protein e
552	6	2.2	140	3	AAO52939	AAO52939 Arabinidops	625	6	2.2	157	2	AAU38867	AAU38867 Neisseria
553	6	2.2	140	3	AAO52944	AAO52944 Arabinidops	626	6	2.2	157	5	ABG68060	ABG68060 Human int
554	6	2.2	140	5	ABP07556	ABP07556 Human CRF	627	6	2.2	158	4	AAO93070	AAO93070 C glutam
555	6	2.2	141	3	AAO18483	AAO18483 Zea may	628	6	2.2	159	4	ABG16754	ABG16754 Novel hum
556	6	2.2	141	6	ABO14990	ABO14990 Human NOV	629	6	2.2	159	5	AAU80548	AAU80548 Human G P
557	6	2.2	141	6	ABM64889	ABM64889 Propionib	630	6	2.2	159	6	ABBA9136	ABBA9136 Phototrab
558	6	2.2	141	7	ADP95085	ADP95085 E. fasciu	631	6	2.2	159	6	ABM69515	ABM69515 Listeria
559	6	2.2	142	4	AAU66857	AAU66857 Propionib	632	6	2.2	160	4	AAU49017	AAU49017 Propionib
560	6	2.2	142	4	ABG08438	ABG08438 Novel hum	633	6	2.2	160	6	ABM45536	ABM45536 Propionib
561	6	2.2	142	6	ABM63376	ABM63376 Propionib	634	6	2.2	161	6	AAAB6305	AAAB6305 Sequence
562	6	2.2	143	3	AAO25805	AAO25805 Arabinidops	635	6	2.2	161	4	AAAB65024	AAAB65024 Human sec
563	6	2.2	143	3	AAO05522	AAO05522 Arabinidops	636	6	2.2	161	5	ABP59105	ABP59105 Protein X
564	6	2.2	144	3	AAO41907	AAO41907 Arabinidops	637	6	2.2	161	5	AAU98692	AAU98692 Arabinidops
565	6	2.2	144	6	ABU48187	ABU48187 Protein e	638	6	2.2	161	5	AAAB6155	AAAB6155 S. pneumo
566	6	2.2	144	6	ABU47025	ABU47025 Protein e	639	6	2.2	163	3	AAO57769	AAO57769 Arabinidops
567	6	2.2	145	4	AAO07941	AAO07941 Human pol	640	6	2.2	163	4	AAO18482	AAO18482 Zea may
568	6	2.2	145	6	ABU31193	ABU31193 Protein e	641	6	2.2	164	4	ABG15828	ABG15828 Novel hum
569	6	2.2	145	6	ABO06440	ABO06440 Altiococc	642	6	2.2	164	5	ABM83046	ABM83046 Translate
570	6	2.2	146	2	AAO59749	AAO59749 Human nor	643	6	2.2	164	5	ABM83046	ABM83046 Translate
571	6	2.2	146	3	AAO25804	AAO25804 Arabinidops	644	6	2.2	165	3	AAO44981	AAO44981 Human mut
572	6	2.2	146	4	AAU49172	AAU49172 Propionib	645	6	2.2	165	3	AAO44980	AAO44980 Human mut
573	6	2.2	146	4	AAU03698	AAU03698 Group B S	646	6	2.2	165	3	AAO44977	AAO44977 Human hyb
574	6	2.2	146	6	ABM45691	ABM45691 Propionib	647	6	2.2	165	4	AAO55041	AAO55041 Arabinidops
575	6	2.2	147	2	AAO98558	AAO98558 H. pylori	648	6	2.2	165	4	AAO55335	AAO55335 Human col
576	6	2.2	147	4	ABO70702	ABO70702 Drosophi	649	6	2.2	165	6	ABM69824	ABM69824 Phototrab
577	6	2.2	147	5	ABP38179	ABP38179 Staphyloc	650	6	2.2	165	6	ABU16597	ABU16597 Protein e
578	6	2.2	147	6	ABU26729	ABU26729 Protein e	651	6	2.2	166	1	AAO30684	AAO30684 Consensus
579	6	2.2	147	6	ABO05506	ABO05506 Altiococc	652	6	2.2	166	1	AAO30686	AAO30686 Consensus
580	6	2.2	148	4	ABG29527	ABG29527 Novel hum	653	6	2.2	166	1	AAO30683	AAO30683 Synthetic
581	6	2.2	148	5	ABM48719	ABM48719 Listeria	654	6	2.2	166	1	AAO40210	AAO40210 Novel int
582	6	2.2	148	5	ABU01559	ABU01559 S. pneumo	655	6	2.2	166	1	AAO50424	AAO50424 Novel int
583	6	2.2	148	6	ABU32695	ABU32695 Protein e	656	6	2.2	166	1	AAO60102	AAO60102 Sequence
584	6	2.2	148	6	ABU40014	ABU40014 Protein e	657	6	2.2	166	1	AAO60101	AAO60101 Sequence
585	6	2.2	148	6	ABU41381	ABU41381 Protein e	658	6	2.2	166	1	AAO60099	AAO60099 Sequence
586	6	2.2	149	4	AAU65789	AAU65789 Propionib	659	6	2.2	166	1	AAO60103	AAO60103 Sequence
587	6	2.2	149	6	ABM62308	ABM62308 Propionib	660	6	2.2	166	1	AAO60104	AAO60104 Sequence
588	6	2.2	149	6	ABM65605	ABM65605 Propionib	661	6	2.2	166	1	AAO60100	AAO60100 Sequence
589	6	2.2	149	6	ABU14997	ABU14997 Protein e	662	6	2.2	166	1	AAO90185	AAO90185 Hybrid al
590	6	2.2	150	4	AAO58822	AAO58822 Human rep	663	6	2.2	166	1	AAO90189	AAO90189 Hybrid al
591	6	2.2	150	4	ABM96353	ABM96353 Human tes	664	6	2.2	166	1	AAO90186	AAO90186 Hybrid al
592	6	2.2	151	2	AAO62721	AAO62721 Streptococ	665	6	2.2	166	1	AAO90190	AAO90190 Hybrid al
593	6	2.2	151	4	ABM02446	ABM02446 Novel hum	666	6	2.2	166	1	AAO90187	AAO90187 Hybrid al
594	6	2.2	151	5	ABM53410	ABM53410 Lactococ	667	6	2.2	166	1	AAO90187	AAO90187 Hybrid al
595	6	2.2	151	6	ABM15935	ABM15935 Mycobacte	668	6	2.2	166	2	AAO20053	AAO20053 Human JFN
596	6	2.2	151	6	ABU52319	ABU52319 Human GPC	669	6	2.2	166	2	AAO42814	AAO42814 Lymphobia
597	6	2.2	151	6	ABU34492	ABU34492 Protein e	670	6	2.2	166	2	AAO42812	AAO42812 Lymphobia
598	6	2.2	151	6	ABU36955	ABU36955 Protein e	671	6	2.2	166	2	AAO67762	AAO67762 Interfero
599	6	2.2	151	7	ADP88421	ADP88421 Ribosomal	672	6	2.2	166	2	AAO70088	AAO70088 Recombina
600	6	2.2	152	6	ABU01766	ABU01766 S. pneumo	673	6	2.2	166	2	AAO50537	AAO50537 Alpha-int
601	6	2.2	152	6	ABM71960	ABM71960 Staphyloc	674	6	2.2	166	2	AAO60141	AAO60141 W. vaccae
602	6	2.2	153	4	AAO78342	AAO78342 Neurolept	675	6	2.2	166	2	AAO14888	AAO14888 Amino aci
603	6	2.2	153	4	AAU62113	AAU62113 Propionib	676	6	2.2	166	3	AAO44975	AAO44975 Human hyb
604	6	2.2	153	4	AAU03673	AAU03673 Group B S	677	6	2.2	166	3	AAO44983	AAO44983 Human mut
605	6	2.2	153	6	ABM58632	ABM58632 Propionib	678	6	2.2	166	3	AAO44974	AAO44974 Human hyb
606	6	2.2	153	6	ABU35768	ABU35768 Protein e	679	6	2.2	166	3	AAO44978	AAO44978 Human hyb
607	6	2.2	153	7	ADP89419	ADP89419 Ribosomal	680	6	2.2	166	3	AAO44978	AAO44978 Human hyb
608	6	2.2	154	3	ABM5023	ABM5023 Human sec	681	6	2.2	166	3	AAO44982	AAO44982 Human mut
609	6	2.2	154	4	AAO78503	AAO78503 Human SP-	682	6	2.2	166	3	AAO28176	AAO28176 Human int

683	6	2.2	166	3	AAB28177	Chimeric	756	6	2.2	167	4	AB559605	AB559605 Drosophila
684	6	2.2	166	3	AAV44834	Mutant by	757	6	2.2	167	4	ABG00676	ABG00676 Novel hum
685	6	2.2	166	3	AAV44832	Mutant by	758	6	2.2	167	6	ABR53739	ABR53739 Protein s
686	6	2.2	166	3	AAV44830	Hybrid in	759	6	2.2	167	6	ABU34732	ABU34732 Protein e
687	6	2.2	166	3	AAV44833	Mutant by	760	6	2.2	168	1	AAE60318	AAE60318 Sequence
688	6	2.2	166	3	AAV44827	Hybrid in	761	6	2.2	168	2	AAE58566	AAE58566 S. Pneumo
689	6	2.2	166	3	AAV44831	Hybrid in	762	6	2.2	168	3	AAE57768	AAE57768 Arbidops
690	6	2.2	166	3	AAV44829	Hybrid in	763	6	2.2	168	4	AAU41736	AAU41736 Propionib
691	6	2.2	166	3	AAV44835	Mutant by	764	6	2.2	168	4	ABG00677	ABG00677 Novel hum
692	6	2.2	166	3	AAV44826	Hybrid in	765	6	2.2	168	6	ABU00544	ABU00544 S. pneumo
693	6	2.2	166	3	AAV44826	Hybrid in	766	6	2.2	168	6	ABM38255	ABM38255 Propionib
694	6	2.2	166	3	AAV44826	Hybrid in	767	6	2.2	169	1	ABU29724	ABU29724 Protein e
695	6	2.2	166	4	AAV49154	Human int	768	6	2.2	170	1	AAE60319	AAE60319 Sequence
696	6	2.2	166	4	AAV49159	Human int	769	6	2.2	170	4	AAE82328	AAE82328 S. epider
697	6	2.2	166	4	AAV49163	Human int	770	6	2.2	170	5	ABU04304	ABU04304 Human col
698	6	2.2	166	4	AAV49168	Human int	771	6	2.2	170	5	ABU04304	ABU04304 Human col
699	6	2.2	166	4	AAV49162	Human int	772	6	2.2	170	5	ABP57955	ABP57955 Transposi
700	6	2.2	166	4	AAV49166	Human int	773	6	2.2	170	6	ABU01903	ABU01903 S. pneumo
701	6	2.2	166	4	AAV49164	Human int	774	6	2.2	170	7	ADC95570	ADC95570 E. faeciu
702	6	2.2	166	4	AAV49164	Human int	775	6	2.2	171	1	AAE60764	AAE60764 Plasmid e
703	6	2.2	166	4	AAV49164	Human int	776	6	2.2	171	6	ABR61643	ABR61643 Human int
704	6	2.2	166	4	AAV49164	Human int	777	6	2.2	171	6	ABR62732	ABR62732 Human alp
705	6	2.2	166	4	AAV49164	Human int	778	6	2.2	172	6	ADA35019	ADA35019 Actinocba
706	6	2.2	166	4	AAV49164	Human int	779	6	2.2	175	6	ABG17565	ABG17565 Novel hum
707	6	2.2	166	4	AAV49164	Human int	780	6	2.2	175	5	AAO15198	AAO15198 Rhodococc
708	6	2.2	166	4	AAV49164	Human int	781	6	2.2	175	5	AAU97143	AAU97143 Rhodococc
709	6	2.2	166	4	AAV49164	Human int	782	6	2.2	175	5	AAE60557	AAE60557 Interfero
710	6	2.2	166	4	AAV49164	Human int	783	6	2.2	176	1	ABJ18773	ABJ18773 pseudomn
711	6	2.2	166	4	AAV49164	Human int	784	6	2.2	176	6	AAE54807	AAE54807 Arbidops
712	6	2.2	166	4	AAV49164	Human int	785	6	2.2	180	2	AAE85436	AAE85436 Streptoc
713	6	2.2	166	4	AAV49164	Human int	786	6	2.2	180	5	ABP69084	ABP69084 Human pol
714	6	2.2	166	4	AAV49164	Human int	787	6	2.2	182	2	AAV21906	AAV21906 Mycobacte
715	6	2.2	166	4	AAV49164	Human int	788	6	2.2	182	2	AAV21906	AAV21906 Amino aci
716	6	2.2	166	4	AAV49164	Human int	789	6	2.2	182	4	AAU17409	AAU17409 Novel sig
717	6	2.2	166	4	AAV49164	Human int	790	6	2.2	182	4	AAU18711	AAU18711 Renal and
718	6	2.2	166	4	AAV49164	Human int	791	6	2.2	182	6	ABU97326	ABU97326 Human pol
719	6	2.2	166	4	AAV49164	Human int	792	6	2.2	182	7	ABU94117	ABU94117 Human nov
720	6	2.2	166	4	AAV49164	Human int	793	6	2.2	182	7	ADC97014	ADC97014 E. faeciu
721	6	2.2	166	4	AAV49164	Human int	794	6	2.2	185	2	AAV34688	AAV34688 Chlamydia
722	6	2.2	166	4	AAV49164	Human int	795	6	2.2	185	3	ABG17774	ABG17774 Arbidops
723	6	2.2	166	4	AAV49164	Human int	796	6	2.2	185	4	ABG01511	ABG01511 Novel hum
724	6	2.2	166	4	AAV49164	Human int	797	6	2.2	186	4	AAE43545	AAE43545 Human pol
725	6	2.2	166	4	AAV49164	Human int	798	6	2.2	186	4	AAE92477	AAE92477 C glutami
726	6	2.2	166	4	AAV49164	Human int	799	6	2.2	187	2	AAE62368	AAE62368 Interfero
727	6	2.2	166	4	AAV49164	Human int	800	6	2.2	187	2	AAE54038	AAE54038 Human pan
728	6	2.2	166	4	AAV49164	Human int	801	6	2.2	188	1	AAE20009	AAE20009 Hybrid hu
729	6	2.2	166	4	AAV49164	Human int	802	6	2.2	188	3	AAE42560	AAE42560 Human ORF
730	6	2.2	166	4	AAV49164	Human int	803	6	2.2	189	1	AAE20108	AAE20108 Sequence
731	6	2.2	166	4	AAV49164	Human int	804	6	2.2	189	1	AAE20104	AAE20104 Sequence
732	6	2.2	166	4	AAV49164	Human int	805	6	2.2	189	1	AAE20105	AAE20105 Sequence
733	6	2.2	166	4	AAV49164	Human int	806	6	2.2	189	1	AAE20111	AAE20111 Sequence
734	6	2.2	166	4	AAV49164	Human int	807	6	2.2	189	1	AAE20112	AAE20112 Sequence
735	6	2.2	166	4	AAV49164	Human int	808	6	2.2	189	1	AAE30228	AAE30228 Sequence
736	6	2.2	166	4	AAV49164	Human int	809	6	2.2	189	1	AAE30179	AAE30179 Sequence
737	6	2.2	166	4	AAV49164	Human int	810	6	2.2	189	1	AAE30183	AAE30183 Sequence
738	6	2.2	166	4	AAV49164	Human int	811	6	2.2	189	1	AAE30164	AAE30164 Sequence
739	6	2.2	166	4	AAV49164	Human int	812	6	2.2	189	1	AAE30165	AAE30165 Sequence
740	6	2.2	166	4	AAV49164	Human int	813	6	2.2	189	1	AAE30101	AAE30101 Sequence
741	6	2.2	166	4	AAV49164	Human int	814	6	2.2	189	1	AAE40741	AAE40741 Sequence
742	6	2.2	166	4	AAV49164	Human int	815	6	2.2	189	1	AAE40179	AAE40179 Sequence
743	6	2.2	166	4	AAV49164	Human int	816	6	2.2	189	1	AAE40123	AAE40123 Sequence
744	6	2.2	166	4	AAV49164	Human int	817	6	2.2	189	1	AAE40127	AAE40127 Sequence
745	6	2.2	166	4	AAV49164	Human int	818	6	2.2	189	1	AAE50306	AAE50306 Human int
746	6	2.2	166	4	AAV49164	Human int	819	6	2.2	189	1	AAE50104	AAE50104 Sequence
747	6	2.2	166	4	AAV49164	Human int	820	6	2.2	189	2	AAE07680	AAE07680 IFN-alpha
748	6	2.2	166	4	AAV49164	Human int	821	6	2.2	189	2	AAE64240	AAE64240 Human int
749	6	2.2	166	4	AAV49164	Human int	822	6	2.2	189	2	AAE62452	AAE62452 Human int
750	6	2.2	166	4	AAV49164	Human int	823	6	2.2	189	2	AAV70373	AAV70373 Human int
751	6	2.2	166	4	AAV49164	Human int	824	6	2.2	189	2	AAE94083	AAE94083 Human int
752	6	2.2	166	4	AAV49164	Human int	825	6	2.2	189	3	AAE59395	AAE59395 Human int
753	6	2.2	166	4	AAV49164	Human int	826	6	2.2	189	3	AAE12867	AAE12867 Human int
754	6	2.2	166	4	AAV49164	Human int	827	6	2.2	189	3	AAE16513	AAE16513 Arbidops
755	6	2.2	166	4	AAV49164	Human int	828	6	2.2	189	4	AAE49779	AAE49779 Human int

829	6	2.2	189	4	AA849780	Abd49780	Human	int	902	6	2.2	206	3	AA645097	Aag45097	Arabidops
830	6	2.2	189	4	AAE14364	Aae14364	Human	int	903	6	2.2	207	4	AA660769	Aab60769	Gene 18 r
831	6	2.2	189	5	AAU84283	Aau84283	Human	end	904	6	2.2	207	4	AA660770	Ab660770	Gene 18 r
832	6	2.2	189	5	ABG68066	ABg68066	Human	int	905	6	2.2	207	6	ABU17138	Abu17138	Protein e
833	6	2.2	189	5	ABG68061	ABg68061	Human	int	906	6	2.2	208	4	AA925077	AAg92507	C glutami
834	6	2.2	189	5	ABG68075	ABg68075	Human	int	907	6	2.2	209	4	AAU34087	Aau34087	Staphyloc
835	6	2.2	189	5	ABG68076	ABg68076	Human	int	908	6	2.2	209	4	AAU36374	Aau36374	Pseudomon
836	6	2.2	189	5	ABG68063	ABg68063	Human	int	909	6	2.2	209	4	AAU36374	Aau36374	Pseudomon
837	6	2.2	189	5	ABG68059	ABg68059	Human	int	910	6	2.2	209	4	AAU36374	Aau36374	Pseudomon
838	6	2.2	189	5	ABG68068	ABg68068	Human	int	911	6	2.2	209	6	AAU36374	Aau36374	Pseudomon
839	6	2.2	189	5	ABG68067	ABg68067	Human	int	912	6	2.2	209	6	AAU36374	Aau36374	Pseudomon
840	6	2.2	189	5	ABG68072	ABg68072	Human	int	913	6	2.2	210	6	ABU27599	Abu27599	Protein e
841	6	2.2	189	5	ABG68071	ABg68071	Human	int	914	6	2.2	211	4	AAU17773	AAu17773	Arabidops
842	6	2.2	189	5	ABG68071	ABg68071	Human	int	915	6	2.2	211	4	AAU17773	AAu17773	Arabidops
843	6	2.2	189	5	ABG68062	ABg68062	Human	int	916	6	2.2	212	6	ABU31486	Abu31486	Human bon
844	6	2.2	189	5	ABG68065	ABg68065	Human	int	917	6	2.2	212	6	ADU34129	Ad34129	Actinotoba
845	6	2.2	189	5	ABG68074	ABg68074	Human	int	918	6	2.2	213	2	AAW23545	AAw23545	Mycoplasma
846	6	2.2	189	5	ABG68076	ABg68076	Human	int	919	6	2.2	213	4	AAU43037	AAu43037	Human ORF
847	6	2.2	189	5	ABG68076	ABg68076	Human	int	920	6	2.2	213	4	ABG00998	ABg00998	Novel hum
848	6	2.2	189	5	ABG68076	ABg68076	Human	int	921	6	2.2	213	6	ABU16388	Abu16388	Protein e
849	6	2.2	189	5	ABG68076	ABg68076	Human	int	922	6	2.2	214	5	ABP26954	ABp26954	Streptoco
850	6	2.2	189	5	ABG68076	ABg68076	Human	int	923	6	2.2	214	5	ABP26954	ABp26954	Streptoco
851	6	2.2	189	5	ABG68076	ABg68076	Human	int	924	6	2.2	215	5	ABP26954	ABp26954	Streptoco
852	6	2.2	189	5	ABG68076	ABg68076	Human	int	925	6	2.2	215	5	ABP26954	ABp26954	Streptoco
853	6	2.2	189	5	ABG68076	ABg68076	Human	int	926	6	2.2	215	5	ABP26954	ABp26954	Streptoco
854	6	2.2	189	5	ABG68076	ABg68076	Human	int	927	6	2.2	216	5	ABP26954	ABp26954	Streptoco
855	6	2.2	189	5	ABG68076	ABg68076	Human	int	928	6	2.2	216	5	ABP26954	ABp26954	Streptoco
856	6	2.2	189	5	ABG68076	ABg68076	Human	int	929	6	2.2	216	6	ABU23076	Abu23076	Protein e
857	6	2.2	189	5	ABG68076	ABg68076	Human	int	930	6	2.2	217	2	AAW23469	AAw23469	Recombina
858	6	2.2	189	5	ABG68076	ABg68076	Human	int	931	6	2.2	217	2	AAW23469	AAw23469	Recombina
859	6	2.2	189	5	ABG68076	ABg68076	Human	int	932	6	2.2	217	2	AAW23469	AAw23469	Recombina
860	6	2.2	189	5	ABG68076	ABg68076	Human	int	933	6	2.2	217	2	AAW23469	AAw23469	Recombina
861	6	2.2	189	5	ABG68076	ABg68076	Human	int	934	6	2.2	217	2	AAW23469	AAw23469	Recombina
862	6	2.2	189	5	ABG68076	ABg68076	Human	int	935	6	2.2	217	2	AAW23469	AAw23469	Recombina
863	6	2.2	189	5	ABG68076	ABg68076	Human	int	936	6	2.2	217	2	AAW23469	AAw23469	Recombina
864	6	2.2	189	5	ABG68076	ABg68076	Human	int	937	6	2.2	217	2	AAW23469	AAw23469	Recombina
865	6	2.2	189	5	ABG68076	ABg68076	Human	int	938	6	2.2	217	2	AAW23469	AAw23469	Recombina
866	6	2.2	189	5	ABG68076	ABg68076	Human	int	939	6	2.2	217	2	AAW23469	AAw23469	Recombina
867	6	2.2	189	5	ABG68076	ABg68076	Human	int	940	6	2.2	217	2	AAW23469	AAw23469	Recombina
868	6	2.2	189	5	ABG68076	ABg68076	Human	int	941	6	2.2	217	2	AAW23469	AAw23469	Recombina
869	6	2.2	189	5	ABG68076	ABg68076	Human	int	942	6	2.2	217	2	AAW23469	AAw23469	Recombina
870	6	2.2	189	5	ABG68076	ABg68076	Human	int	943	6	2.2	217	2	AAW23469	AAw23469	Recombina
871	6	2.2	189	5	ABG68076	ABg68076	Human	int	944	6	2.2	217	2	AAW23469	AAw23469	Recombina
872	6	2.2	189	5	ABG68076	ABg68076	Human	int	945	6	2.2	218	3	AAU39423	AAu39423	Arabidops
873	6	2.2	189	5	ABG68076	ABg68076	Human	int	946	6	2.2	218	6	ADU12321	Ad12321	Alloioococ
874	6	2.2	189	5	ABG68076	ABg68076	Human	int	947	6	2.2	219	2	AAU39423	AAu39423	Arabidops
875	6	2.2	189	5	ABG68076	ABg68076	Human	int	948	6	2.2	219	2	AAU39423	AAu39423	Arabidops
876	6	2.2	189	5	ABG68076	ABg68076	Human	int	949	6	2.2	219	2	AAU39423	AAu39423	Arabidops
877	6	2.2	189	5	ABG68076	ABg68076	Human	int	950	6	2.2	219	4	AAU39423	AAu39423	Arabidops
878	6	2.2	189	5	ABG68076	ABg68076	Human	int	951	6	2.2	220	4	AAU39423	AAu39423	Arabidops
879	6	2.2	189	5	ABG68076	ABg68076	Human	int	952	6	2.2	220	6	AAU39423	AAu39423	Arabidops
880	6	2.2	189	5	ABG68076	ABg68076	Human	int	953	6	2.2	221	2	AAU39423	AAu39423	Arabidops
881	6	2.2	189	5	ABG68076	ABg68076	Human	int	954	6	2.2	221	2	AAU39423	AAu39423	Arabidops
882	6	2.2	189	5	ABG68076	ABg68076	Human	int	955	6	2.2	221	3	AAU39423	AAu39423	Arabidops
883	6	2.2	189	5	ABG68076	ABg68076	Human	int	956	6	2.2	221	3	AAU39423	AAu39423	Arabidops
884	6	2.2	189	5	ABG68076	ABg68076	Human	int	957	6	2.2	221	5	AAU39423	AAu39423	Arabidops
885	6	2.2	189	5	ABG68076	ABg68076	Human	int	958	6	2.2	221	7	AAU39423	AAu39423	Arabidops
886	6	2.2	189	5	ABG68076	ABg68076	Human	int	959	6	2.2	222	4	AAU39423	AAu39423	Arabidops
887	6	2.2	189	5	ABG68076	ABg68076	Human	int	960	6	2.2	222	6	AAU39423	AAu39423	Arabidops
888	6	2.2	189	5	ABG68076	ABg68076	Human	int	961	6	2.2	223	3	AAU39423	AAu39423	Arabidops
889	6	2.2	189	5	ABG68076	ABg68076	Human	int	962	6	2.2	223	4	AAU39423	AAu39423	Arabidops
890	6	2.2	189	5	ABG68076	ABg68076	Human	int	963	6	2.2	223	5	AAU39423	AAu39423	Arabidops
891	6	2.2	189	5	ABG68076	ABg68076	Human	int	964	6	2.2	223	5	AAU39423	AAu39423	Arabidops
892	6	2.2	189	5	ABG68076	ABg68076	Human	int	965	6	2.2	224	5	AAU39423	AAu39423	Arabidops
893	6	2.2	189	5	ABG68076	ABg68076	Human	int	966	6	2.2	225	4	AAU39423	AAu39423	Arabidops
894	6	2.2	189	5	ABG68076	ABg68076	Human	int	967	6	2.2	226	4	AAU39423	AAu39423	Arabidops
895	6	2.2	189	5	ABG68076	ABg68076	Human	int	968	6	2.2	227	6	AAU39423	AAu39423	Arabidops
896	6	2.2	189	5	ABG68076	ABg68076	Human	int	969	6	2.2	228	3	AAU39423	AAu39423	Arabidops
897	6	2.2	189	5	ABG68076	ABg68076	Human	int	970	6	2.2	228	3	AAU39423	AAu39423	Arabidops
898	6	2.2	189	5	ABG68076	ABg68076	Human	int	971	6	2.2	228	3	AAU39423	AAu39423	Arabidops
899	6	2.2	189	5	ABG68076	ABg68076	Human	int	972	6	2.2	228	3	AAU39423	AAu39423	Arabidops
900	6	2.2	189	5	ABG68076	ABg68076	Human	int	973	6	2.2	228	3	AAU39423	AAu39423	Arabidops
901	6	2.2	189	5	ABG68076	ABg68076	Human	int	974	6	2.2	228	3	AAU39423	AAu39423	Arabidops

Abu22153 Protein e
Aag52948 Arabidops
Abm69350 Photorhab
Ada36158 Kinetoba
Aau33381 Novel hum
Aag91949 C glutam
Aag75438 Human col
Abg26456 Novel hum
Aau4978 Hepatitis
Abb61621 Drosophila
Aar12363 HTLV-1 en
Aar42265 Decorin s
Aag49817 Arabidops
Abp28742 Streptoco
Abu46706 Protein e
Abu3508 Protein e
Aag92550 C glutam
Abg15597 Novel hum
Aay59839 Human nor
Aay74002 Human pro
Aag46706 Arabidops
Aag19425 Arabidops
Aag19438 Arabidops
Aag46702 Arabidops
Aag19578 Arabidops
Aag46689 Arabidops

ALIGNMENTS

RESULT 1
AAB20104 standard; protein; 276 AA.

XX AAB20104;
XX
XX 23-APR-2001 (first entry)
XX
XX Moraxella catarrhalis BASB11 protein.
XX
XX BASB11; infection; otitis media; pneumonia; diagnosis; therapy;
XX
XX antibacterial; antimicrobial; vaccine.
XX
XX Moraxella catarrhalis.
XX
XX MO200100837-A1.
XX
XX 04-JAN-2001.
XX
XX 23-JUN-2000; 2000WO-EP005852.
XX
XX 25-JUN-1999; 99GB-00014945.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Thonard J;
XX
XX WPI; 2001-123013/13.
XX
XX N-PSDB; AAF30040.
XX
XX New BASB11 polypeptides of Moraxella catarrhalis useful for diagnostic,
XX
XX prophylactic and therapeutic purposes against microbial diseases,
XX
XX preferably bacterial infections.
XX
XX Claim 1; Page 63; 79pp; English.
XX
XX The present sequence is that of BASB11 protein from Moraxella
XX
XX catarrhalis strain MC2931 (ATCC 43617), a causative agent of otitis media
XX
XX in children and pneumonia in adults. The invention provides BASB11
XX
XX polypeptides, and polynucleotides encoding them, as well as expression
XX
XX vectors, host cells and methods for producing BASB11 polypeptides using
XX
XX recombinant methods. Also claimed is a vaccine composition comprising a

CC BASB11 polypeptide, an immunogenic fragment of a BASB11 polypeptide, or
CC a polypeptide having at least 85% amino acid sequence identity to
CC BASB11, or comprising a polynucleotide encoding such a polypeptide. A
CC claimed method of diagnosing a Moraxella infection involves identifying a
CC BASB11 polypeptide or antibody. A claimed therapeutic composition useful
CC in treating humans with M. catarrhalis infection comprises at least 1
CC antibody directed against a BASB11 polypeptide. BASB11 polypeptides
CC also have utility in raising specific antibodies, and in screening for
CC antibacterial drugs

XX SQ Sequence 276 AA;

Query Match 100.0%; Score 276; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 9, 3e-251; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;

QY 1 MNPFGKINGICALAGSIALAGCSNOSNEPAISKTAAQTIKYGVAGCEQAVAEVAGVAK 60
DB 1 MNPFGKINGICALAGSIALAGCSNOSNEPAISKTAAQTIKYGVAGCEQAVAEVAGVAK 60
QY 61 EKNVITVELVFNPDYAMPNSAVSGEILDANMOKPYLEKDSQSKGNNIVITVNTVYP 120
DB 61 EKNVITVELVFNPDYAMPNSAVSGEILDANMOKPYLEKDSQSKGNNIVITVNTVYP 120
QY 121 LAGYSTKIKTLNELKDGATIVPNDPSNLARALILLEKQGLIKDKNTNLFSTLIDIVEN 180
DB 121 LAGYSTKIKTLNELKDGATIVPNDPSNLARALILLEKQGLIKDKNTNLFSTLIDIVEN 180
QY 181 PKGLVKEVDTSVVARAIDVDVLA VNNNNAAGVGLTASENGVEFVEKDSPTVNIIVARA 240
DB 181 PKGLVKEVDTSVVARAIDVDVLA VNNNNAAGVGLTASENGVEFVEKDSPTVNIIVARA 240
QY 241 DNKOSKATQDFVKAQYQDVEAEAKKQFKQGVINGW 276
DB 241 DNKOSKATQDFVKAQYQDVEAEAKKQFKQGVINGW 276

RESULT 2

AA017582
ID AA017582 standard; protein; 276 AA.

XX AA017582;
XX
XX 19-JUL-2002 (first entry)
XX
XX M catarrhalis MCA101896 protein SEQ ID NO: 44.
XX
XX Moraxella; vaccine; respiratory tract infection; anti-inflammatory;
XX
XX auditory; antibacterial; otitis media; sinusitis; pneumonia.
XX
XX Moraxella catarrhalis.
XX
XX MO200218595-A2.
XX
XX 07-MAR-2002.
XX
XX 28-AUG-2001; 2001WO-CA001221.
XX
XX 28-AUG-2000; 2000US-0228294P.
XX
XX 28-AUG-2000; 2000US-0228295P.
XX
XX 28-AUG-2000; 2000US-0228296P.
XX
XX 29-AUG-2000; 2000US-0228438P.
XX
XX 29-AUG-2000; 2000US-0228439P.
XX
XX 29-AUG-2000; 2000US-0228440P.
XX
XX 29-AUG-2000; 2000US-0228441P.
XX
XX 29-AUG-2000; 2000US-0228442P.
XX
XX 29-AUG-2000; 2000US-0228443P.
XX
XX 29-AUG-2000; 2000US-0228511P.
XX
XX 29-AUG-2000; 2000US-0228512P.
XX
XX 29-AUG-2000; 2000US-0228742P.
XX
XX 29-AUG-2000; 2000US-0228743P.
XX
XX 01-SEP-2000; 2000US-0229465P.
XX
XX 01-SEP-2000; 2000US-0229474P.

PR 01-SEP-2000; 2000US-0229475P.
 PR 01-SEP-2000; 2000US-0229478P.
 PR 05-SEP-2000; 2000US-0229740P.
 PR 05-SEP-2000; 2000US-0229803P.
 PR 05-SEP-2000; 2000US-0229804P.
 PR 05-SEP-2000; 2000US-0229805P.
 PR 05-SEP-2000; 2000US-0229806P.
 PR 05-SEP-2000; 2000US-0229809P.
 PR 05-SEP-2000; 2000US-0229811P.
 PR 06-SEP-2000; 2000US-0230214P.
 PR 06-SEP-2000; 2000US-0230250P.
 PR 06-SEP-2000; 2000US-0230252P.
 PA (AVET) AVENTIS PASTEUR LTD.
 PI Loomore S, Wang J, Bradley B, Ochs M, Yang Y;
 XX WPI; 2002-404555/43.
 XX N-PSDB; AAL46514.
 DR
 XX
 PT Moraxella polypeptide and polynucleotides useful as vaccine for
 PT immunizing a host e.g. humans against disease e.g. otitis media,
 PT pneumonia, caused by infection of the bacteria.
 XX
 PS Claim 28; Fig 43; 277pp; English.
 XX
 CC The present invention provides the protein and coding sequences of
 CC proteins from Moraxella catarrhalis. These can be used to produce
 CC vaccines which protect against M. catarrhalis infection, which can cause
 CC otitis media, respiratory infection, sinusitis, and pneumonia. The
 CC present sequence is a protein of the invention
 XX
 SQ Sequence 276 AA;
 Query Match 88.0%; Score 243; DB 5; Length 276;
 Best Local Similarity 100.0%; Pred. No. 9.4e-220;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 DB 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 QY 94 HKPYLEKDSQEKGLNLTIVGNFVYPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 DB 94 HKPYLEKDSQEKGLNLTIVGNFVYPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 QY 154 ILEKQGLIKLKONTNLFSTTLDIVENPKKLVKEVDTSVAARALDDVDLAVVNNNYAQ 213
 DB 154 ILEKQGLIKLKONTNLFSTTLDIVENPKKLVKEVDTSVAARALDDVDLAVVNNNYAQ 213
 QY 214 VGLTASENVFVDEKSPYVNIIVARADKSKALIDPFYKATQOTDEVEAEAKKQFDDGYI 273
 DB 214 VGLTASENVFVDEKSPYVNIIVARADKSKALIDPFYKATQOTDEVEAEAKKQFDDGYI 273
 QY 274 KGW 276
 DB 274 KGW 276
 RESULT 3
 ABUS3462
 ID ABUS3462 standard; protein; 276 AA.
 AC ABUS3462;
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #20989.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS Moraxella catarrhalis.
 XX

PN MO20027183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342821P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELITR) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA39332.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 63386; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_ptc_sequences
 XX
 SQ Sequence 276 AA;
 Query Match 88.0%; Score 243; DB 6; Length 276;
 Best Local Similarity 100.0%; Pred. No. 9.4e-220;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 DB 34 TAAQTIRKGVAGPEQVAEVAQVAKKYNLTVELVEFNDYAMPNSAVSKGELDANAMQ 93
 QY 94 HKPYLEKDSQEKGLNLTIVGNFVYPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 DB 94 HKPYLEKDSQEKGLNLTIVGNFVYPLAGYSTKIKTLNELKDGATIAVNDPSNLARAL 153
 QY 154 ILEKQGLIKLKONTNLFSTTLDIVENPKKLVKEVDTSVAARALDDVDLAVVNNNYAQ 213

Db 154 ILLEKQGLIKLKNNTNFTTLDIVENPKLVKEVDTSVAARAIDVLAIVNNVIAQ 213
 QY 214 VGLTASNGVFEVDKSPYNIIVARADNKSALIDPFVKAQYOTDEVEAEAKKQFDGYI 273
 Db 214 VGLTASNGVFEVDKSPYNIIVARADNKSALIDPFVKAQYOTDEVEAEAKKQFDGYI 273
 QY 274 KGW 276
 Db 274 KGW 276

RESULT 4
 AAO17813
 ID AAO17813 standard; protein: 272 AA.

AAO17813;
 AC AAO17813;
 XX 05-AUG-2002 (first entry)
 DT H influenzae BVH-NTNH12 protein SEQ ID NO: 24.
 DE XX
 XX Haemophilus influenzae infection; BVH-NTNH1; otitis media; BVH-NTNH2;
 KW sinusitis; bronchitis; pneumonia; meningitis; bacteraemia; BVH-NTNH3;
 KW BVH-NTNH4; BVH-NTNH5; BVH-NTNH6; BVH-NTNH7; BVH-NTNH8; BVH-NTNH9;
 KW BVH-NTNH10; BVH-NTNH11; BVH-NTNH12; antiinflammatory; auditory;
 KW antibacterial; vaccine.
 XX Haemophilus influenzae.
 OS WO200228889-A2.
 XX 11-APR-2002.
 PF 02-OCT-2001; 2001WO-CA001402.
 XX 02-OCT-2000; 2000US-0236712P.
 PR 02-OCT-2000; 2000US-0236712P.
 XX (SHIR-) SHIRE BIOCHEM INC.
 PA Hamel J, Couture F, Brodeur BR, Martin D, Ouellet C, Tremblay M,
 PI Charbonneau A, Vaysialer C;
 DR WPI; 2002-435325/46.
 DR N-PSDB; AAL46811.
 XX Novel isolated Haemophilus influenzae polypeptides BVH-NTNH1-12, useful
 PT for inducing protective immune responses against H. influenzae in animals
 PT and for treating otitis media, sinusitis, bronchitis and pneumonia.
 XX Claim 17; Fig 24; 58pp; English.
 XX The present invention provides the protein and coding sequences of
 CC Haemophilus influenzae BVH-NTNH1-12. The sequences can be used in the
 CC production of a vaccine to protect against, and in the diagnosis of, H.
 CC influenzae infection, which can lead to otitis media, sinusitis,
 CC bronchitis, pneumonia, meningitis and bacteraemia. The present sequence
 CC is a protein of the invention
 XX
 SQ Sequence 272 AA;

Query Match 6.9%; Score 19; DB 5; Length 272;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 AAO35474
 ID AAO35474 standard; protein: 273 AA.

XX AAO35474;
 AC 14-FEB-2002 (first entry)
 DT Haemophilus influenzae cellular proliferation protein #115.
 DE Haemophilus influenzae cellular proliferation protein; antibiotic;
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 OS Haemophilus influenzae.
 XX
 XX WO200170955-A2.
 XX 27-SEP-2001.
 PD 21-MAR-2001; 2001WO-US009180.
 PF 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS53333.

PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 11067; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 273 AA;

Query Match 6.9%; Score 19; DB 4; Length 273;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
 ABB76073
 ID ABB76073 standard; protein: 273 AA.

AC ABB76073;
 XX 15-JUL-2002 (first entry)
 XX Haemophilus influenzae BASB202 protein.
 DE BASB202; NTHi; infection; vaccine; genetic immunisation; auditory;
 KM antiinflammatory; antibacterial; immunostimulant; otitis media.
 XX Haemophilus influenzae.
 XX Key Location/Qualifiers
 FT Region 1..15 /note= "potential T-helper cell epitope"
 FT Region 21..28 /note= "potential B-cell epitope"
 FT Region 34..42 /note= "potential T-helper cell epitope"
 FT Region 43..47 /note= "potential B-cell epitope"
 FT Region 56..60 /note= "potential B-cell epitope"
 FT Region 66..74 /note= "potential T-helper cell epitope"
 FT Region 69..74 /note= "potential B-cell epitope"
 FT Region 78..84 /note= "potential B-cell epitope"
 FT Region 89..97 /note= "potential T-helper cell epitope"
 FT Region 93..104 /note= "potential B-cell epitope"
 FT Region 107..126 /note= "potential T-helper cell epitope"
 FT Region 121..127 /note= "potential B-cell epitope"
 FT Region 129..134 /note= "potential B-cell epitope"
 FT Region 141..148 /note= "potential B-cell epitope"
 FT Region 150..169 /note= "potential T-helper cell epitope"
 FT Region 176..182 /note= "potential B-cell epitope"
 FT Region 181..189 /note= "potential T-helper cell epitope"
 FT Region 200..208 /note= "potential T-helper cell epitope"
 FT Region 224..228 /note= "potential B-cell epitope"
 FT Region 229..240 /note= "potential T-helper cell epitope"
 FT Region 236..244 /note= "potential B-cell epitope"
 FT Region 249..256 /note= "potential B-cell epitope"
 FT Region 259..267 /note= "potential B-cell epitope"
 FT Region /note= "potential B-cell epitope"
 XX WO200224729-A2.
 XX 28-MAR-2002.
 XX 18-SEP-2001; 2001WO-BP010979.
 XX 19-SEP-2000; 2000GB-00022392.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Thonnard J;
 XX WPI; 2002-383180/41.
 DR N-PDB; ABL56953, ABL56954, ABL56956.
 DR

XX New isolated BASB202 polypeptide of nontypeable Haemophilus influenzae
 PT useful for diagnosing a disease and in generating an immune response in
 PT an animal.
 XX Claim 3; Page 75; 90pp; English.
 PS
 XX The present sequence is the protein sequence for the BASB202 protein of
 CC nontypeable Haemophilus influenzae (NTHi) strain 3224A (ATCC PT-1816)
 CC isolated from an otitis media patient in the USA, NTHi strain 3219C
 CC isolated from an otitis media patient in the USA, NTHi strain 810956
 CC isolated from a meningitis patient in the Netherlands. Variability analysis of
 CC BASB202 protein among NTHi strains revealed 99-100% sequence identity.
 CC The invention provides BASB202 polypeptides and polynucleotides, vectors,
 CC host cells, and methods for producing the polypeptides by recombinant
 CC methods. Claimed vaccine compositions comprise a BASB202 polypeptide or
 CC polynucleotide. A claimed method of diagnosing NTHi infection involves
 CC identifying a BASB202 polypeptide or an antibody that is immunospecific
 CC for the polypeptide. A claimed therapeutic composition useful for
 CC treating humans with NTHi disease comprises an antibody directed against
 CC a BASB202 polypeptide. NTHi diseases include otitis media, pneumonia,
 CC sinusitis, nosocomial infections, invasive disease, chronic otitis media
 CC with hearing loss, fluid accumulation in the middle ear, auditory nerve
 CC damage, delayed speech learning, infection of the upper respiratory
 CC tract, and inflammation of the middle ear. B-cell epitopes and T-helper
 CC cell epitopes from BASB202 may also be useful in vaccine compositions
 XX
 SQ Sequence 273 AA;
 Query Match 6.9%; Score 19; DB 5; Length 273;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 107 LNNLYVGNTPYFYPAGYS 125
 DB 104 LNNLYVGNTPYFYPAGYS 122
 RESULT 7
 ABB76074
 ID ABB76074 standard; protein; 273 AA.
 XX ABB76074;
 AC
 XX 15-JUL-2002 (first entry)
 DT
 XX Haemophilus influenzae BASB202 protein.
 DE
 XX BASB202; NTHi; infection; vaccine; genetic immunisation; auditory;
 KM antiinflammatory; antibacterial; immunostimulant; otitis media.
 XX Haemophilus influenzae.
 OS
 XX Key Location/Qualifiers
 FT Region 1..15 /note= "potential T-helper cell epitope"
 FT Region 21..28 /note= "potential B-cell epitope"
 FT Region 34..42 /note= "potential T-helper cell epitope"
 FT Region 43..47 /note= "potential B-cell epitope"
 FT Region 56..60 /note= "potential B-cell epitope"
 FT Region 66..74 /note= "potential T-helper cell epitope"
 FT Region 69..74 /note= "potential B-cell epitope"
 FT Region 78..84 /note= "potential B-cell epitope"
 FT Region 89..97 /note= "potential T-helper cell epitope"
 FT Region /note= "potential B-cell epitope"
 FT

CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

CC Sequence 273 AA;

Query Match 6.9%; Score 19; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 LNNLVIGNTFVYPLAGYS 125
104 LNNLVIGNTFVYPLAGYS 122

RESULT 9
ABU40467
ID ABU40467 standard; protein; 271 AA.

XX ABB76094;

DT 15-JUL-2002 (first entry)

DE Haemophilus influenzae BASB202 protein T-helper cell epitope.

KM BASB202; NTH1; infection; vaccine; genetic immunisation; auditory;
KM anti-inflammation; antibacterial; immunostimulant; otitis media;
KM T lymphocyte; epitope.

OS Haemophilus influenzae.

PN WO200224729-A2.

PD 28-MAR-2002.

PF 18-SEP-2001; 2001WO-EP010979.

PR 19-SEP-2000; 2000GB-00022992.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Thonard J;

DR WPI; 2002-383180/41.

PT New isolated BASB202 polypeptide of nontypeable Haemophilus influenzae
PT useful for diagnosing a disease and in generating an immune response in
PT an animal.

PS Example 13; Page 72; 90pp; English.

CC The present sequence is a potential T-helper cell epitope of the BASB202
CC protein of nontypeable Haemophilus influenzae (NTH1). It corresponds to
CC amino acid residues 107-126 of the BASB202 sequences given in ABB76073
CC and ABB76074. T-helper cell epitopes are peptides bound to human
CC leukocyte antigen (HLA) class II molecules and recognised by T-helper
CC cells. Peptides comprising BASB202 T-cell epitopes (see ABB76090-99) can
CC be useful (preferably when conjugated to a peptide, polypeptide or
CC polysaccharide) in a vaccine composition for the prevention of NTH1-
CC related diseases, such as otitis media, pneumonia, sinusitis, nosocomial
CC infections, invasive disease, chronic otitis media with hearing loss,
CC fluid accumulation in the middle ear, auditory nerve damage, delayed
CC speech learning, infection of the upper respiratory tract, and
CC inflammation of the middle ear

XX Sequence 20 AA;

Query Match 5.8%; Score 16; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 LVIVGNTFVYPLAGYS 125
DB 1 LVIVGNTFVYPLAGYS 16

RESULT 10
ABU40467
ID ABU40467 standard; protein; 271 AA.

XX ABU40467;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #25994.

KM Antisense, prokaryotic essential gene; cell proliferation; drug design.

OS Proteus sp.

PN WO20027183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR N-PSDB; ACA44337.

PF 18-SEP-2001; 2001WO-EP010979.

PR 19-SEP-2000; 2000GB-00022992.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Thonard J;

DR WPI; 2002-383180/41.

PT New isolated BASB202 polypeptide of nontypeable Haemophilus influenzae

PT useful for diagnosing a disease and in generating an immune response in

PT an animal.

PS Example 13; Page 72; 90pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than S. aureus, S. typhimurium,

CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 270 AA;
 Query Match 4.7%; Score 13; DB 6; Length 270;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 134 LKDGATIAVPNDP 146
 130 LKDGATIAVPNDP 142
 Db
 RESULT 13
 ABU39463
 ID ABU39463 standard; protein; 276 AA.
 XX
 AC ABU39463;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #24990.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Pasteurella multocida.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (BLIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GF, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR N-PDB; AC433333.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 67387; 1766pp; English.

CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 276 AA;
 Query Match 4.3%; Score 12; DB 6; Length 276;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 155 LLEKQGLIKKD 166
 155 LLEKQGLIKKD 166
 Db
 RESULT 14
 AAB70412
 ID AAB70412 standard; protein; 67 AA.
 XX
 AC AAB70412;
 XX
 DT 02-MAY-2001 (first entry)
 XX
 DE A. actinomycetemcomitans immunoreactive protein SEQ ID NO:15.
 XX
 KW Actinobacillus actinomycetemcomitans; microbial; infection; vaccine;
 KW identification; localised juvenile periodontitis; antibacterial;
 KW antiinflammatory.
 XX
 OS Actinobacillus actinomycetemcomitans.
 XX
 PN WO200111081-A2.
 XX
 PD 15-FEB-2001.
 XX
 PF 04-AUG-2000; 2000WO-US021340.
 PF 06-AUG-1999; 99US-0147551P.
 PR
 XX
 PA (IVIG-) IVIGENE CORP.
 PI Prognulske-Fox A, Handfield M, Brady LJ, Hillman JD;
 XX
 DR WPI: 2001-202779/20.
 DR N-PDB; AAF59422.
 XX
 PT Identifying microbial polynucleotides, useful for vaccine design,
 PT diagnostics and antibiotherapy, comprises isolating clones of a microbe's
 PT expression library reactive with antibodies against microbe proteins
 PT produced during in vivo growth.
 XX
 PS Example 3; Page 65; 68pp; English.

The present invention describes a method (M1) for identifying a
 CC polynucleotide (I) of a microbe (M) that is expressed in vivo. (M1)
 CC comprises: (a) absorbing antibodies (Ab) against antigens that are
 CC expressed by (M) in vivo and in vitro with cells or cellular extracts of
 CC (M) that have been grown in vitro; (b) isolating unadsorbed Abs; and (c)
 CC probing an expression library of (M)'s DNA/RNA with unadsorbed Abs, where
 CC (I) that is expressed in vivo is identified. The method can be used for
 CC identifying antigens expressed during an actual microbial infection. The
 CC identified polynucleotides are useful for vaccine design, diagnostics and
 CC antibiotherapy, in particular for the diagnosis and therapy of
 CC Actinobacillus actinomycetemcomitans infection, which is the etiologic
 CC agent for localised juvenile periodontitis. The present sequence

CC represents an *Actinobacillus actinomycetemcomitans* immunoreactive
 CC protein, which is used in an example from the present invention
 CC
 XX
 SQ Sequence 67 AA:

Query Match 4.0%; Score 11; DB 4; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.027; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 TIKGVWAGPE 48
 19 TIKGVWAGPE 29

RESULT 15
 ABU27902
 ID ABU27902 standard; protein; 271 AA.

AC ABU27902;
 DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #13429.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Enterobacter cloacae.
 OS WO200277183-A2.

XX PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.

ER (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.

DR N-PSDB; ACJ31772.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 55826; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; (7) identifying a compound that influences the activity of
 CC proliferation; (8) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcr_sequences

SQ Sequence 271 AA;

Query Match 4.0%; Score 11; DB 6; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.088; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKDSPPYN 234
 219 FVEDKDSPPYN 229

RESULT 16
 ABU48271
 ID ABU48271 standard; protein; 271 AA.

AC ABU48271;
 DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #33798.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX *Salmonella typhi*.
 OS WO200277183-A2.

XX PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.

ER (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.

DR N-PSDB; ACJ52141.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 76195; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 271 AA;

Query Match Best Local Similarity 4.0%; Score 11; DB 6; Length 271;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 FVEDKSDSPYVN 224
|||||
DB 219 FVEDKSDSPYVN 229

RESULT 17

ABU16868 ID ABU16868 standard; protein; 276 AA.

XX AC ABU16868;

DT 19-JUN-2003 (first entry)

XX DE Protein encoded by prokaryotic essential gene #2395.

XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Acinetobacter baumannii.

XX PN W0200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002MO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;

XX PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;

XX DR N-PSDB; ACA20738.

XX DR WPI; 2003-029926/02.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PT Claim 25; SEQ ID NO 44792; 1766gp; English.

XX The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 276 AA;

Query Match Best Local Similarity 3.6%; Score 10; DB 6; Length 276;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 ARAIDVDVLA 204
|||||
DB 190 ARAIDVDVLA 199

RESULT 18

ADA35415 ID ADA35415 standard; protein; 286 AA.

XX AC ADA35415;

DT 20-NOV-2003 (first entry)

XX DE Acinetobacter baumannii protein #2576.

XX KM Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;

XX KW plant biocontrol agent.

XX OS Acinetobacter baumannii.

XX PN US6562958-B1.

XX PD 13-MAY-2003.

XX PF 04-JUN-1999; 99US-00328352.

XX PR 09-JUN-1998; 98US-0086701P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton G, Bush D;

XX DR WPI; 2003-576092/54.

XX DR N-PSDB; ADA31289.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX plants.
XX Example; SEQ ID NO 6702; 328bp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 286 AA;
XX
Query Match 3.6%; Score 10; DB 6; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 195 ARAIDVDVLA 204
Db 200 ARAIDVDVLA 209
XX
RESULT 19
ABR53692
ID ABR53692 standard; protein; 125 AA.
XX
AC ABR53692;
XX
XX 20-JUN-2003 (first entry)
XX
DE Protein sequence #SEQ ID 2249.
XX
XX
XX Multiprotein complex; eukaryote; drug target; diagnosis.
XX
XX Saccharomyces cerevisiae.
XX
XX EPI258494-A1.
XX
XX 20-NOV-2002.
XX
XX 20-DEC-2001; 2001EP-00130253.
XX
XX 15-MAY-2001; 2001EP-00111774.
XX
XX (CELL-) CELLZONE AG.
XX
XX Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
XX Marzloch M, Schultz JD, Superti-Furga GD;
XX
XX WPI: 2003-250078/25.
XX N-PSDB; ACC61734.
XX
XX New isolated protein complexes useful for diagnosing a disease or
XX disorder, or as a target for an active agent of a pharmaceutical,
XX preferably a drug target in the treatment or prevention of disease or
XX disorder.
XX
XX Disclosure; SEQ ID NO 2249; 17pp + Sequence listing; English.
XX
XX The invention relates to multiprotein complexes from eukaryotes. Proteins
XX of the invention and DNA sequences encoding them are given in records
XX ABR53692-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
XX obtainable by using a protein as a bait and isolating the set of proteins
XX which is attached thereto from cells. Such protein complexes may comprise
XX up to 30 distinct proteins. Protein complexes of the invention are useful
XX for diagnosing a disease or disorder, or as a target for an active agent

CC of a pharmaceutical, preferably a drug target in the treatment or
CC prevention of a disease or disorder. Note: The sequence data for this
CC patent is not represented in the printed specification, but is based on
CC sequence information supplied by the European Patent Office. The complete
CC document is available on CD-ROM
XX
SQ Sequence 125 AA;
XX
Query Match 3.3%; Score 9; DB 6; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 132 NELKDGATI 140
Db 74 NELKDGATI 82
XX
RESULT 20
ABU16864
ID ABU16864 standard; protein; 241 AA.
XX
AC ABU16864;
XX
XX 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #2391.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Acinetobacter baumannii.
XX
XX OS
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX N-PSDB; ACR20734.
XX
XX WPI: 2003-029926/02.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX
XX
XX Claim 25; SEQ ID NO 44788; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway;
XX the gene product or that has an activity against a biological pathway;
XX (8) identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs; or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-regulated gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs; or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match 3.3%; Score 9; DB 6; Length 241;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 256 AA;

RESULT 21
ABU33203
ID ABU33203 standard; protein; 256 AA.

Query Match 3.3%; Score 9; DB 6; Length 256;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ABU33203;
19-JUN-2003 (first entry)
Protein encoded by Prokaryotic essential gene #18730.

140 IAVPNDPSN 148
125 IAVPNDPSN 133

Antisense; prokaryotic essential gene; cell proliferation; drug design.
Legionella pneumophila.

ABU23447
ID ABU23447 standard; protein; 262 AA.

WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WC-US009107.

ABU23447;
19-JUN-2003 (first entry)
Protein encoded by Prokaryotic essential gene #8974.

21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.

Antisense; prokaryotic essential gene; cell proliferation; drug design.
Bordetella pertussis.
WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WC-US009107.

(ELIT-) ELITRA PHARM INC.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
Wali D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
N-PSDB; ACA37073.
WPI; 2003-029926/02.

21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to

Claim 25; SEQ ID NO 61127; 1766bp; English.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
Wali D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated

for homologous nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 51371; 1766bp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 262 AA;
 Query Match 3.3%; Score 9; DB 6; Length 262;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 133 ELKDGATIA 141
 |||||
 Db 119 ELKDGATIA 127

RESULT 23
 ABU27880
 ID ABU27880 standard; protein; 270 AA.
 XX
 XX ABU27880;
 AC
 XX 19-JUN-2003 (first entry)
 DT
 XX
 DE Protein encoded by Prokaryotic essential gene #13407.
 XX
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Enterobacter cloacae.
 OS
 XX WO200277183-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WC-US009107.
 PF
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX

PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR MPI: 2003-029926/02.
 DR N-PSDB; AC311750.
 XX
 PT New antisense nucleic acid, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 55804; 1766bp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 270 AA;
 Query Match 3.3%; Score 9; DB 6; Length 270;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 176 DIVENPKXL 184
 |||||
 Db 165 DIVENPKXL 173

RESULT 24
 ABM67576
 ID ABM67576 standard; protein; 272 AA.
 XX
 XX ABM67576;
 AC
 XX 20-NOV-2003 (first entry)
 DT
 XX
 DE Photorhabdus luminescens protein sequence #673.
 XX
 XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 OS
 XX Photorhabdus luminescens.

XX MO200294867-A2.
 XX 28-NOV-2002.
 XX 07-FEB-2002; 2002MO-IB003040.
 XX 07-FEB-2001; 2001FR-00001659.
 XX (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX Duchaud E, Taourit S, Glaeser P, Frangeul L, Kunst F, Danchin A,
 XX Buchrieser C;
 XX WPI; 2003-148459/14.
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX Claim 2; SEQ ID NO 673; 1205bp; French.
 XX The invention relates to the isolation of genes and their encoded
 XX proteins from Photorhabdus luminescens. The isolated sequences are
 XX sources of probes and primers for detecting the genome of P. luminescens
 XX and related species; to study polymorphisms; for gene analysis and for
 XX detection/amplification of the genes. Antibodies (Ab) raised against the
 XX polypeptides encoded by the genes are used for detection/identification
 XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 XX carry a gene-containing vector are used to select compounds that
 XX modulate, regulate, induce or inhibit expression of the genes in plants,
 XX animals or microorganisms other than P. luminescens and are able to alter
 XX response or sensitivity to toxins and antibiotics produced by P.
 XX recombinant production of the proteins, particularly toxins and
 XX antibacterials useful as insecticides, bactericides and fungicides. The
 XX genes, proteins, vectors containing the genes and Ab are also useful
 XX therapeutically (to treat microbial infection by bacteria or fungi that
 XX are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 XX biopesticides. Other uses of the genes and the proteins are as virulence
 XX factors and for identifying targets of human diseases for which P.
 XX luminescens is a model (particularly plague and whooping cough). This
 XX sequence represents one of the isolated P. luminescens proteins
 XX
 XX Sequence 272 AA;
 XX
 XX Query Match 3.3%; Score 9; DB 6; Length 272;
 XX Best Local Similarity 100.0%; Pred. No. 6.6;
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 158 KQGLIKLKD 166
 XX |||||
 XX Db 151 KQGLIKLKD 159
 XX
 XX RESULT 25
 XX ABB48120
 XX ID ABB48120 standard; protein; 273 AA.
 XX
 XX ABB48120;
 XX
 XX 05-FEB-2002 (first entry)
 XX
 XX Listeria monocytogenes protein #824.
 XX
 XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 XX vitamin B12; bacterial infection; disease.
 XX
 XX Listeria monocytogenes.
 XX
 XX WO200177335-A2.
 XX
 XX 18-OCT-2001.

XX 11-APR-2001; 2001MO-FR001118.
 XX 11-APR-2000; 2000FR-00004629.
 XX (INSP) INST PASTEUR.
 XX Buchrieser C, Frangeul L, Couve E, Ruenick C, Faint H, Dehoux P,
 XX Dussange O, Chetoui F, Nedjati H, Glaeser P, Kunst F, Cossart P,
 XX Daniels U, Goebel W, Kreft U, Kuhn M, Ng E, Vazquez-Boland JA,
 XX Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
 XX Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L,
 XX Perez-Diaz U, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 XX Madueno E, De Pablos B, Weiland U, Kaerst U, Entian K, Hauf U,
 XX Rose M, Voss H;
 XX WPI; 2002-010914/01.
 XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 XX and prevention of Listeria and related bacterial infections, and related
 XX polypeptides.
 XX Claim 6; SEQ ID NO 825; 192bp; French.
 XX The present invention relates to the genome sequence of Listeria
 XX monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 XX it are useful for selecting probes and primers for detecting genes in L.
 XX monocytogenes and related organisms, and for studying genetic
 XX polymorphisms and other genomes. The present sequence is a protein
 XX encoded by the genome sequence of the present invention. Proteins
 XX expressed from the genome sequence are useful for raising specific
 XX antibodies. Identification of L. monocytogenes and related organisms, and
 XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 XX B12. The genome sequence and proteins encoded by it are also useful for
 XX selecting compounds that regulate gene expression and cell replication
 XX and modulate L. monocytogenes-related diseases. In addition, the genome
 XX sequence and proteins encoded by it are useful in pharmaceutical and
 XX vaccines compositions for the treatment or prevention of infections by L.
 XX monocytogenes and related organisms. Note: The sequence data for this
 XX patent did not form part of the printed specification, but was obtained
 XX in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pcc_sequences
 XX
 XX Sequence 273 AA;
 XX
 XX Query Match 3.3%; Score 9; DB 5; Length 273;
 XX Best Local Similarity 100.0%; Pred. No. 6.7;
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 140 IAVPNDPSN 148
 XX |||||
 XX Db 136 IAVPNDPSN 144
 XX
 XX RESULT 26
 XX ABB32634
 XX ID ABB32634 standard; protein; 273 AA.
 XX
 XX ABB32634;
 XX
 XX 19-JUN-2003 (first entry)
 XX
 XX Protein encoded by Prokaryotic essential gene #18161.
 XX
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 XX Listeria monocytogenes.
 XX
 XX WO200277183-A2.
 XX
 XX 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002MO-US009107.

RESULT 28
 ADC97607
 ID ADC97607 standard; protein: 291 AA.
 XX
 AC ADC97607;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE E. faecium protein sequence SEQ ID 7234.
 XX
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KM abdominal-pelvic infection.
 XX
 OS Enterococcus faecium.
 XX
 PN US6583275-B1.
 XX
 PD 24-JUN-2003.
 XX
 PF 30-JUN-1998; 98US-00107532.
 XX
 PR 02-JUL-1997; 97US-0051571P.
 XX
 PR 14-MAY-1998; 98US-0085598P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI, 2003-799836/75.
 DR
 DR N-PSDB; ADC93953.
 XX
 PT New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.
 PS
 PS Example 1; SEQ ID NO 7234; 243pp; English.
 XX
 CC The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising) 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridizing to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acid is useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of Candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.
 CC
 SQ Sequence 291 AA;
 Query Match 3.3%; Score 9; DB 7; Length 291;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 231 PYNIVAR 239
 DB 240 PYNIVAR 248
 RESULT 29
 ADA35637
 ID ADA35637 standard; protein: 296 AA.
 XX

AC ADA35637;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #2798.
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KM plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI, 2003-576092/54.
 DR
 DR N-PSDB; ADA31511.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 PS
 PS Example; SEQ ID NO 6924; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 CC
 SQ Sequence 296 AA;
 Query Match 3.3%; Score 9; DB 6; Length 296;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 148 NLARALLIL 156
 DB 164 NLARALLIL 172
 RESULT 30
 ID AAU18017
 ID AAU18017 standard; protein: 96 AA.
 XX
 AC AAU18017;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immunoglobulin polypeptide SEQ ID NO 162.
 XX
 KW Immunoglobulin; signal transduction pathway protein; cancer;
 KW antisense therapy; gene therapy; neurological disorder; renal disorder;
 KW cardiovascular disorder; gastrointestinal disorder; pulmonary disorder;
 KW reproductive disorder; immune system disorder; proliferative disorder;
 KM muscular disorder.
 XX
 OS Homo sapiens.
 OS
 PN WO200155315-A2.

XX	02-AUG-2001.	
PD		
XX	17-JUN-2001.	2001MO-US001326
PF		
XX	31-JUN-2000.	2000US-0179065P
FR	04-FEB-2000.	2000US-0180668P
FR	24-FEB-2000.	2000US-0184664P
FR	02-MAR-2000.	2000US-0186350P
FR	16-MAR-2000.	2000US-0189874P
PR	17-MAR-2000.	2000US-0190076P
PR	18-MAR-2000.	2000US-0198132P
PR	19-MAR-2000.	2000US-0205515P
PR	07-JUN-2000.	2000US-0209647P
PR	28-JUN-2000.	2000US-0214866P
PR	30-JUN-2000.	2000US-0215135P
PR	07-JUL-2000.	2000US-0216677P
PR	11-JUL-2000.	2000US-0216880P
PR	11-JUL-2000.	2000US-0217487P
PR	14-JUL-2000.	2000US-0219590P
PR	26-JUL-2000.	2000US-0220563P
PR	26-JUL-2000.	2000US-0220646P
PR	14-AUG-2000.	2000US-0224518P
PR	14-AUG-2000.	2000US-0224519P
PR	14-AUG-2000.	2000US-0224519P
PR	14-AUG-2000.	2000US-0225213P
PR	14-AUG-2000.	2000US-0225214P
PR	14-AUG-2000.	2000US-0225566P
PR	14-AUG-2000.	2000US-0225566P
PR	14-AUG-2000.	2000US-0225668P
PR	14-AUG-2000.	2000US-0225668P
PR	14-AUG-2000.	2000US-0225677P
PR	14-AUG-2000.	2000US-0225757P
PR	14-AUG-2000.	2000US-0225758P
PR	14-AUG-2000.	2000US-0225758P
PR	18-AUG-2000.	2000US-0225759P
PR	18-AUG-2000.	2000US-0226279P
PR	22-AUG-2000.	2000US-0226668P
PR	22-AUG-2000.	2000US-0226668P
PR	22-AUG-2000.	2000US-0227182P
PR	30-AUG-2000.	2000US-0227682P
PR	30-AUG-2000.	2000US-0228094P
PR	01-SEP-2000.	2000US-0229287P
PR	01-SEP-2000.	2000US-0229287P
PR	01-SEP-2000.	2000US-02292943P
PR	01-SEP-2000.	2000US-02292943P
PR	01-SEP-2000.	2000US-0229345P
PR	05-SEP-2000.	2000US-0229509P
PR	05-SEP-2000.	2000US-0229513P
PR	06-SEP-2000.	2000US-0230437P
PR	06-SEP-2000.	2000US-0230438P
PR	08-SEP-2000.	2000US-0231142P
PR	08-SEP-2000.	2000US-0231143P
PR	08-SEP-2000.	2000US-0231144P
PR	08-SEP-2000.	2000US-0231143P
PR	08-SEP-2000.	2000US-0231143P
PR	08-SEP-2000.	2000US-0232080P
FR	12-SEP-2000.	2000US-0232081P
FR	12-SEP-2000.	2000US-0232081P
PR	14-SEP-2000.	2000US-0232397P
PR	14-SEP-2000.	2000US-0232398P
PR	14-SEP-2000.	2000US-0232399P
PR	14-SEP-2000.	2000US-0232401P
PR	14-SEP-2000.	2000US-0232401P
PR	14-SEP-2000.	2000US-0233063P
PR	14-SEP-2000.	2000US-0233064P
PR	14-SEP-2000.	2000US-0233065P
PR	21-SEP-2000.	2000US-0234223P
PR	21-SEP-2000.	2000US-0234774P
PR	25-SEP-2000.	2000US-0234977P
PR	25-SEP-2000.	2000US-0234988P
PR	26-SEP-2000.	2000US-0235484P
PR	27-SEP-2000.	2000US-0235834P
PR	27-SEP-2000.	2000US-0236327P
PR	29-SEP-2000.	2000US-0236537P

DR MPI: 2001-45725/49.
DR N-PSDB: AAS28805.
XX Isolated novel immunoglobulin polypeptide for monitoring the presence and
PT progression of diseases and for diagnosis.
XX Claim 11; SEQ ID NO 162; 551bp; English.
XX Sequences AAU17977-AAU18087 represent immunoglobulin polypeptides of the
CC invention. The polypeptides and their associated polynucleotides can be
CC used to diagnose a pathological condition or a susceptibility to a
CC pathological condition in a subject by determining the presence or
CC absence of a mutation in a DNA sequence or determining the presence or
CC amount of expression of the protein. Alternatively the identification of
CC a binding partner to a sequence allows determination of changes in
CC protein activity. The sequences can be used as research tools for
CC receptors or other signal transduction pathway proteins that interact
CC with the polypeptides of the invention and can be used to treat, prevent
CC or diagnose various types of disorders such as neurological disorders,
CC cardiovascular disorders, gastrointestinal disorders, reproductive
CC disorders, immune system disorders, renal disorders, muscular disorders,
CC pulmonary disorders, proliferative disorders and cancer. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 96 AA;
Query Match 2.9%; Score 8; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 PSNLRAL 153
Db 76 PSNLRAL 83
RESULT 31
ABBI0497
ID ABBI0497 standard; protein, 96 AA.
XX
AC ABBI0497;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human CDNA SEQ ID NO: 805.
XX
XX Human; gene therapy; neural disorder; immune system disorder;
KM muscular disorder; reproductive disorder; gastrointestinal disorder;
KM pulmonary disorder; cardiovascular disorder; renal disorder;
KM proliferative disorder; inflammation.
XX
OS Homo sapiens.
XX
PV WO200154474-A2.
XX
FD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001349.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184654P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0235135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-OUL-2000; 2000US-0216890P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251993P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen Ca, Barash SC, Ruben SW;
PI
XX WPI, 2001-476161/51.
DR
XX N-PSDB; ABA06719.
XX
XX Isolated nucleic acid molecule encoding an inflammation-associated
PT polypeptide is used in preventing, treating or ameliorating a medical
PT condition.
XX
XX
PS Claim 11: SEQ ID NO 805; 859bp + Sequence Listing; English.
XX
XX
XX The present invention provides human cDNAs, proteins and related genomic
CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a protein of the invention
XX
XX
SO Sequence 96 AA;

Query Match 2.9%; Score 8; DB 4; Length 96;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 PSNLRAL 153
DB 76 PSNLRAL 83
RESULT 32
ID ABBE7084 standard; protein; 96 AA.
XX
AC ABBE7084;
XX
DT 09-DEC-2002 (first entry)
XX
XX Human polypeptide SEQ ID NO 805.
DE
XX Human; nootropic; neuroprotective; cytoskeletal; dermatological; virucide;
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; virucide;
KW antiparkinsonian; antitickling; antianemic; antidiabetic; cancer;
KW antithrombotic; hepatoprotective; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephroprotective; gene therapy; vaccine.
OS Homo sapiens.
XX
XX US2002090672-A1.
XX
PD 11-JUL-2002.
XX
PF 17-JAN-2001; 2001US-00764853.
XX
XX 31-JAN-2000; 2000US-01790655P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217486P.
PR 14-JUL-2000; 2000US-0218280P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236372P.
PR 29-SEP-2000; 2000US-0236373P.
PR 29-SEP-2000; 2000US-0236387P.
PR 29-SEP-2000; 2000US-0236388P.
PR 29-SEP-2000; 2000US-0236399P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.

02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 01-NOV-2000; 2000US-0244809P.
PR 17-NOV-2000; 2000US-0248299P.
PR 08-DEC-2000; 2000US-0251836P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX
XX (ROSE/) ROSEN C A.
PA (RUBEN/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2002-681727/73.
DR N-PSDB; ABV84056.
XX
XX Novel polypeptide useful for diagnosis, prognosis, prevention, and
PT treatment of immune, hyperproliferative, renal, respiratory,
FT cardiovascular, reproductive, endocrine, gastrointestinal and
PT neurological disorders.
XX
XX
XX Claim 11; SEQ ID NO 805; 369pp + Sequence Listing; English.
PS
XX The invention relates to novel genes (ABV83682-ABV84101) and proteins
CC (ABP66710-ABP67129) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breasts, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischemias; (d) wound healing
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 96 AA;
SQ
Query Match 2.9%; Score 8; DB 5; Length 96;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 PSNLRAL 153
Db 76 PSNLRAL 83
RESULT 33
ADBJ1641
ID ADBJ1641 standard; protein; 96 AA.
XX
XX ADBJ1641;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Human novel protein SEQ ID NO 162.
XX
XX Gene therapy; human; immunoglobulin; cancer; lung cancer; leukaemia;
KM ovarian cancer; epithelial cancer; hyperplasia; Gaucher's disease; AIDS;
KM arrhythmia; cardiac oedema; ischaemia; pneumonia; cystic fibrosis;
KM asthma; sarcoidosis; rhinitis; anaemia; inflammation; sinusitis;
XX chronic obstructive pulmonary disease; infectious disease.
XX

OS Homo sapiens.
XX
XX US2003077606-A1.
XX
XX 24-APR-2003.
XX
XX 07-MAR-2002; 2002US-00091438.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226881P.
PR 22-AUG-2000; 2000US-0226886P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR

PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 29-SEP-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239393P.
 PR 13-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 05-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-0076487P.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-615993/58.
 DR N-PSDB; ADB31530.
 XX
 PT New human immunoglobulin superfamily of polypeptide and genes, useful for
 PT treating, preventing or diagnosing e.g. cancers (lung cancer, leukemia,
 PT ovarian epithelial cancer, etc.), hyperplasia, ischemia, pneumonia or
 PT AIDS.
 XX
 PS Claim 12; SEQ ID NO 162; 213pp; English.
 XX
 CC The invention relates to an isolated polypeptide, which comprises the
 CC human immunoglobulin superfamily of proteins. The polypeptide or
 CC polynucleotide is useful for treating, preventing or ameliorating a
 CC medical condition e.g. cancers (lung cancer, leukemia, ovarian
 CC epithelial cancer, etc.), hyperplasia, Gaucher's disease, AIDS,
 CC arrhythmia, cardiac oedema, ischaemia, pneumonia, cystic fibrosis,
 CC asthma, sarcoidosis, rinitis, anaemia, leukaemia, inflammatory
 CC sinusitis, chronic obstructive pulmonary disease, infectious diseases
 CC etc. The polypeptide or polynucleotide is also useful for diagnosing any
 CC of these diseases or a susceptibility to the disease. The present
 CC sequence represent the amino acid sequence of a novel human protein.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format direct from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=20030077606
 CC
 XX
 SQ Sequence 96 AA;
 XX
 QY 146 PSNLARAL 153
 DB 76 PSNLARAL 83
 XX
 RESULT 34
 ID AAO05643 standard, protein; 152 AA.
 AC AAO05643;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 19535.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004927.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR N-PSDB; AAI85574.
 XX

Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukemia, inflammation and immune disorders.

Claim 20; SEQ ID NO 19535; 1339pp + Sequence Listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and the encoded proteins (AA000010-AA013910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Notes: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pot_sequences](http://wipo.int/pub/published_pot_sequences)

prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein. Notes: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pot_sequences](http://wipo.int/pub/published_pot_sequences)

Query Match 2.9%; Score 8; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 2.9%; Score 8; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 146 PSNTARAL 153
Db 116 PSNTARAL 123

Qy 113 VGNTFYVP 120
Db 75 VGNTFYVP 82

RESULT 35

AAU36063 standard; protein; 240 AA.

AAU36063 standard; protein; 240 AA.

AAU36063;

14-FEB-2002 (first entry)

Klebsiella pneumoniae cellular proliferation protein #51.

Antisense; prokaryotic cellular proliferation protein; antibiotic;
antibacterial; drug design.

Klebsiella pneumoniae.

MO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US009180.

21-MAR-2000; 2000US-0191078P.

23-MAY-2000; 2000US-0206848P.

26-MAY-2000; 2000US-0207127P.

23-OCT-2000; 2000US-0242578P.

27-NOV-2000; 2000US-0253625P.

22-DEC-2000; 2000US-0257931P.

16-FEB-2001; 2001US-0269308P.

(ELIT-) ELITRA PHARM INC.

Haseelbeck R, Ohlsen KL, Zyckind JW, Wall D, Trawick JD, Carr GT,

Yamamoto RT, Xu HH;

WPI; 2001-611495/70.

N-PSDB; AAS53922.

New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids.

Example 3; SEQ ID NO 11656; 511pp; English.

The invention relates to antisense inhibitors of genes essential to

RESULT 36

ABP26478 standard; protein; 263 AA.

ABP26478 standard; protein; 263 AA.

ABP26478;

02-JUN-2002 (first entry)

Streptococcus pyogenes polypeptide SEQ ID NO 2132.

Streptococcus; GAS; GBS; group B streptococcus; *Streptococcus agalactiae*;

group A streptococcus; *Streptococcus pyogenes*; antibacterial;

antiinflammatory; infection; vaccine; meningitis; gene therapy.

Streptococcus pyogenes.

MO200234771-A2.

02-MAY-2002.

29-OCT-2001; 2001WO-GB004789.

27-OCT-2000; 2000GB-00026333.

24-NOV-2000; 2000GB-00028727.

07-MAR-2001; 2001GB-00005640.

(CHIR-) CHIRON SPA.

(GENO-) INST GENOMIC RES.

Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;

Tetcelin H;

WPI; 2002-352536/38.

N-PSDB; ABN57109.

New Streptococcus protein for the treatment or prevention of infection or

disease caused by Streptococcus bacteria, such as meningitis, and for

detecting a compound that binds to the protein.

Claim 1; Page 3365; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B

Streptococcus/GAS (*Streptococcus agalactiae*) or group A streptococcus/GAS

CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (1) Nucleic acids encoding (1), ABN6044-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 263 AA;
XX
QY Query Match 2.9%; Score 8; DB 5; Length 263;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 253 KAYOTDEV 260
240 KAYOTDEV 247
XX
RESULT 37
ABU25052
ID ABU25052 standard; protein; 263 AA.
XX
AC ABU25052;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #10579.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Clostridium difficile.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
N-PSDB; ACA28922.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 52976; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 263 AA;
XX
QY Query Match 2.9%; Score 8; DB 6; Length 263;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 229 DSPYVNI 236
216 DSPYVNI 223
XX
DB 216 DSPYVNI 223
XX
RESULT 38
ABU23277
ID ABU23277 standard; protein; 265 AA.
XX
AC ABU23277;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #8804.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bordetella pertussis.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
N-PSDB; ACA27147.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 51201; 1766pp; English.

PS The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway of
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 265 AA;

Query Match 2.9%; Score 8; DB 6; Length 265;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 131 LNELEKGA 138
Db 120 LNELEKGA 127

RESULT 39
ADC96816
ID ADC96816 standard; protein; 269 AA.
XX
AC ADC96816;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 6443.
XX
KM Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
XX abdominal-pelvic infection.
XX Enterococcus faecium.
OS
PN US6583275-B1.
XX
PD 24-JUN-2003.
XX
PF 30-JUN-1998; 98US-00107532.
XX
PR 02-JUL-1997; 97US-0051571P.
XX 14-MAY-1998; 98US-0085598P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;

XX WPI; 2003-799836/75.
DR N-PSDB; ADC93162.
XX

PT New isolated nucleic acid derived from *Enterococcus faecium* encoding an
XX nucleic acid polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
XX infection.

PS Example 1; SEQ ID NO 6443; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from
CC *Enterococcus faecium* encoding an *Enterococcus faecium* polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridizing to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids are useful for diagnosing pathological conditions
CC resulting from *E. faecium* bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of *Candida albicans* -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating *Enterococcus faecium* infections. The present sequence represents
XX one of the disclosed *E. faecium* proteins.

SQ Sequence 269 AA;

Query Match 2.9%; Score 8; DB 7; Length 269;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 ELKDGATT 140
Db 118 ELKDGATT 125

RESULT 40
ABU21415
ID ABU21415 standard; protein; 270 AA.
XX
AC ABU21415;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #6942.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Burkholderia fungorum.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009-107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948893.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362689P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW,
XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.

DR N-PSDB; ACA25285.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 45339; 1766bp; English.

XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 270 AA;

Query Match

Best Local Similarity 2.9%; Score 8; DB 6; Length 270;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 AVNDPSN 148

Db 132 AVNDPSN 139

Search completed: June 16, 2004, 11:15:51
Job time : 90 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:13:13 ; Search time 21 Seconds
(without alignments)

1264.231 Million cell updates/sec

Title: US-10-018-672-2

Sequence: 1 MNFGKINGICALAGALAG.....TDEVEAKKQKQGVKGM 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : PIR 78:*

- 1: Dir1:*
- 2: Dir2:*
- 3: Dir3:*
- 4: Dir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	12.3	277	2 JN0751	Outer membrane 30K
2	19	6.9	273	2 B64082	outer membrane pro
3	12	4.3	263	2 JN0753	outer membrane 29.
4	11	4.0	271	2 G90653	probable lipoprote
5	11	4.0	271	2 G85504	probable lipoprote
6	11	4.0	271	2 A10532	probable lipoprote
7	7	3.3	125	2 A48107	DNA-directed RNA p
8	9	3.3	273	2 A11110	probable lipoprote
9	9	3.3	273	2 A82266	lipoprotein yacC V
10	9	3.3	278	2 A13551	ABC transporter su
11	9	3.3	391	1 MHUBT	ig mu heavy chain
12	8	2.9	74	2 B81284	hypothetical prote
13	8	2.9	89	2 H75053	molybdopterin conv
14	8	2.9	259	2 AB3108	outer membrane lip
15	8	2.9	259	2 G98178	probable periplasm
16	8	2.9	271	2 E64744	probable lipoprote
17	8	2.9	446	2 A83213	probable ATP-depen
18	7	2.5	87	2 C81400	protein P46F5.13 (
19	7	2.5	104	2 B88029	hypothetical prote
20	7	2.5	114	2 G90820	hypothetical prote
21	7	2.5	114	2 G90851	hypothetical prote
22	7	2.5	114	2 C90974	hypothetical prote
23	7	2.5	114	2 C90902	hypothetical prote
24	7	2.5	114	2 D90911	hypothetical prote
25	7	2.5	114	2 F90971	hypothetical prote
26	7	2.5	114	2 H91065	hypothetical prote
27	7	2.5	114	2 D85679	unknown protein en
28	7	2.5	115	2 F85819	unknown protein en
29	7	2.5	115	2 H85627	hypothetical prote

30	7	2.5	123	2 F85744	unknown protein en
31	7	2.5	124	2 H64608	hypothetical prote
32	7	2.5	126	2 AB2559	integrational host f
33	7	2.5	137	2 A64312	hypothetical prote
34	7	2.5	141	2 UB3359	plasmacyte sprea
35	7	2.5	154	2 S74334	biotin carboxyl ca
36	7	2.5	156	2 I67751	dopamine receptor
37	7	2.5	188	2 C97438	hypothetical prote
38	7	2.5	189	2 AF2656	conserved hypothet
39	7	2.5	193	2 J02185	coat protein - app
40	7	2.5	196	2 AC3629	3-octaprenyl-4-hyd
41	7	2.5	207	2 B53801	chitin synthase (E
42	7	2.5	215	2 T03680	plasma membrane pr
43	7	2.5	232	2 T03678	mrRB protein precu
44	7	2.5	233	1 C39142	hypothetical prote
45	7	2.5	234	2 B71905	hypothetical prote
46	7	2.5	239	2 AB1301	probable anion-upt
47	7	2.5	241	2 D87547	3-oxoadipate CoA-t
48	7	2.5	247	2 F81410	hydrogenase isoenz
49	7	2.5	259	2 H64211	guanylate kinase (
50	7	2.5	260	2 B82957	probable TonB-depe
51	7	2.5	261	2 G82750	outer membrane pro
52	7	2.5	262	2 B82170	peptide ABC transp
53	7	2.5	268	2 J01473	pancreatic elastas
54	7	2.5	268	2 S35589	peptide transport
55	7	2.5	268	2 B64877	peptide transport
56	7	2.5	268	2 C90862	hypothetical prote
57	7	2.5	268	2 F85756	hypothetical prote
58	7	2.5	268	2 AC0656	peptide transport
59	7	2.5	269	2 B64134	peptide transport
60	7	2.5	271	2 A71802	probable outer mem
61	7	2.5	271	2 D64715	outer membrane pro
62	7	2.5	271	2 AG0131	probable lipoprote
63	7	2.5	271	2 AH0287	peptide transport
64	7	2.5	273	2 AB1472	probable lipoprote
65	7	2.5	274	2 AH0160	probable exported
66	7	2.5	276	2 JN0752	outer membrane 30.
67	7	2.5	276	2 AC1746	conserved lipoprot
68	7	2.5	276	2 A11376	conserved lipoprot
69	7	2.5	286	2 F86664	outer membrane lip
70	7	2.5	286	2 G86664	outer membrane lip
71	7	2.5	297	2 F64470	sulfate permease (
72	7	2.5	298	1 XWBO	ADP,ATP carrier pr
73	7	2.5	298	1 A29132	ADP,ATP carrier pr
74	7	2.5	298	1 S03894	ADP,ATP carrier pr
75	7	2.5	298	2 S31814	ADP,ATP carrier pr
76	7	2.5	298	2 B43646	ADP,ATP carrier pr
77	7	2.5	298	2 AH3116	conserved hypothet
78	7	2.5	300	2 T15206	hypothetical prote
79	7	2.5	300	2 B75619	probable cobalamu
80	7	2.5	300	2 AC2831	transcription regu
81	7	2.5	300	2 C97609	transcription regu
82	7	2.5	306	2 G86168	hypothetical 27.0K
83	7	2.5	307	2 F66703	conserved hypothet
84	7	2.5	311	2 AB1915	hydroxymethylbilan
85	7	2.5	311	2 B81188	hydroxymethylbilan
86	7	2.5	314	2 T46743	carbamate kinase (
87	7	2.5	334	2 T20728	hypothetical prote
88	7	2.5	341	2 F75334	3-oxoacyl-acyl car
89	7	2.5	344	2 G70458	hypothetical prote
90	7	2.5	347	2 G00006	haemoglobin - biac
91	7	2.5	350	2 S39156	outer membrane pro
92	7	2.5	352	2 F64751	probable ABC-type
93	7	2.5	352	2 A85531	hypothetical prote
94	7	2.5	352	2 B90680	hypothetical prote
95	7	2.5	380	2 B82634	threonine dehydrat
96	7	2.5	386	2 S75400	hypothetical prote
97	7	2.5	397	2 B87343	conserved hypothet
98	7	2.5	400	2 C82503	hypothetical prote
99	7	2.5	404	2 T36254	probable valine-py
100	7	2.5	406	2 AD0833	glucose membrane
101	7	2.5	412	1 S74961	glucose dehydrogen
102	7	2.5	419	2 G70652	probable sets prot

103	7	2.5	420	2	E70914	probable lipo prot	176	6	2.2	79	2	H42505	K-ORF-B protein -
104	7	2.5	432	2	A12613	hypothetical prote	177	6	2.2	82	2	A98836	hypothetical prote
105	7	2.5	432	2	G97395	hypothetical prote	178	6	2.2	82	2	C85694	unknown protein en
106	7	2.5	434	2	G75339	conserved hypotet	179	6	2.2	84	2	AB1622	ribosomal protein
107	7	2.5	454	2	D69066	amidase - Melanoc	180	6	2.2	84	2	AB1259	ribosomal protein
108	7	2.5	457	2	I51660	dopamine D1B recep	181	6	2.2	85	2	G81150	hypothetical prote
109	7	2.5	468	2	T21331	hypothetical prote	182	6	2.2	90	2	F91254	hypothetical prote
110	7	2.5	471	2	S48265	hypothetical prote	183	6	2.2	90	2	D70678	hypothetical prote
111	7	2.5	473	2	H72419	virulence factor M	184	6	2.2	90	2	B86095	hypothetical prote
112	7	2.5	475	2	A41271	dopamine receptor	185	6	2.2	90	2	F65209	hypothetical 10.5
113	7	2.5	477	2	DYHND5	dopamine receptor	186	6	2.2	92	2	AF1012	conserved hypotet
114	7	2.5	486	2	B55886	dopamine receptor	187	6	2.2	92	2	B75559	ribosomal protein
115	7	2.5	491	2	UC5312	UTP-hexose-1-phosp	188	6	2.2	92	2	A47120	alcohol dehydrogen
116	7	2.5	530	2	D82063	probable thiamin A	189	6	2.2	92	2	T20113	hypothetical prote
117	7	2.5	538	2	G72539	probable CTP synth	190	6	2.2	93	2	CG9141	hypothetical prote
118	7	2.5	549	2	H96944	arginine degradati	191	6	2.2	95	1	H18P14	internal protein I
119	7	2.5	553	2	E89725	protein C02C6.2 [1	192	6	2.2	96	2	S48006	IPI protein - phag
120	7	2.5	554	2	T18858	hypothetical prote	193	6	2.2	97	2	D41892	hydrogenase HupF -
121	7	2.5	572	2	T37128	hypothetical prote	194	6	2.2	98	2	D83872	hypothetical prote
122	7	2.5	575	2	B64174	hypothetical prote	195	6	2.2	99	2	D83452	utase operon 23X
123	7	2.5	580	2	S76668	hypothetical prote	196	6	2.2	99	2	AE0926	probable lipoprote
124	7	2.5	600	2	A45112	major paraflogella	197	6	2.2	100	2	F64225	ribosomal protein
125	7	2.5	610	2	T47725	hypothetical prote	198	6	2.2	102	2	A53308	hypothetical prote
126	7	2.5	610	2	T16194	hypothetical prote	199	6	2.2	104	2	AC1159	hypothetical prote
127	7	2.5	611	2	F75095	probable asparagin	200	6	2.2	104	2	AC1518	hypothetical prote
128	7	2.5	616	2	H82130	proteinase IV VCI3	201	6	2.2	106	1	Z2BPT9	gene 49.2 protein
129	7	2.5	652	2	C95129	DNA ligase, NAD-de	202	6	2.2	106	2	T17814	hypothetical prote
130	7	2.5	652	2	H97959	DNA ligase (NAD)	203	6	2.2	107	2	F53306	K1e8 protein - pla
131	7	2.5	670	2	B70145	periplasmic protei	204	6	2.2	109	2	D70701	hypothetical prote
132	7	2.5	691	2	C64548	outer membrane pro	205	6	2.2	109	2	AG3602	hydrolypoamide d
133	7	2.5	694	2	E72272	conserved hypotet	206	6	2.2	110	2	G84101	flagellar protein
134	7	2.5	696	2	E95938	probable ATP-bind	207	6	2.2	111	2	T00668	hypothetical prote
135	7	2.5	783	1	A48984	DNA excision repa	208	6	2.2	112	2	G69781	thioredoxin homolo
136	7	2.5	852	2	H81906	probable lipo-pro	209	6	2.2	113	2	H75183	hypothetical prote
137	7	2.5	852	2	B81110	protein-P11 uridy	210	6	2.2	115	2	SS9099	ribosomal protein
138	7	2.5	889	2	AE1270	DNA polymerase I	211	6	2.2	116	2	A83640	hypothetical prote
139	7	2.5	895	2	UC6015	chitin synthase I	212	6	2.2	117	2	G70306	ribosomal protein
140	7	2.5	905	2	AH2892	hypothetical prote	213	6	2.2	118	2	AB3036	hypothetical prote
141	7	2.5	905	2	C97668	ABC transporter re	214	6	2.2	121	2	AB1872	hypothetical prote
142	7	2.5	1036	2	A82357	probable multidrug	215	6	2.2	121	2	AB1872	ribosomal protein
143	7	2.5	1036	2	S14032	kinesin-related pr	216	6	2.2	122	1	S33697	hypothetical prote
144	7	2.5	1095	2	E96744	probable oligopep	217	6	2.2	124	2	H89794	hypothetical prote
145	7	2.5	1118	2	H97298	subtilisin like pr	218	6	2.2	125	2	H47677	hypothetical prote
146	7	2.5	1130	2	T20288	hypothetical prote	219	6	2.2	126	2	AD2313	hypothetical prote
147	7	2.5	1239	2	G02750	DNA-directed DNA p	220	6	2.2	128	2	T14487	vary hypotetrical
148	7	2.5	1333	2	S38635	blastocyst polyprot	221	6	2.2	128	2	H97101	uncharacterized pr
149	7	2.5	1795	2	F97713	190K antigen precu	222	6	2.2	128	2	H97354	uncharacterized pr
150	7	2.5	1994	2	D86452	protein F6N18.13	223	6	2.2	128	2	AF3006	acetyltransferase
151	7	2.5	2154	2	F83068	hypothetical prote	224	6	2.2	128	2	F88277	kasugamycin acetyl
152	7	2.5	2154	2	F82885	hypothetical prote	225	6	2.2	129	2	I56195	gene Tap-1 protein
153	6	2.2	18	2	G42753	interferon alpha (226	6	2.2	130	2	PC4423	transcription regu
154	6	2.2	21	2	S57202	vitronectin-bindin	227	6	2.2	133	2	T24039	hypothetical prote
155	6	2.2	22	2	A61413	interferon alpha (228	6	2.2	134	2	B83011	phosphoribosyl-AMP
156	6	2.2	22	2	B61413	interferon alpha (229	6	2.2	134	2	D39039	allergen Amb a I -
157	6	2.2	22	2	S09063	proteasome chain 2	230	6	2.2	134	2	D64381	conserved hypotet
158	6	2.2	33	2	I68894	gene TAP1 protein	231	6	2.2	135	2	D61243	hypothetical prote
159	6	2.2	33	2	D42753	interferon alpha (232	6	2.2	136	2	T68814	hypothetical prote
160	6	2.2	49	2	S55324	endo-beta-1,6-gluc	233	6	2.2	137	2	A70909	hypothetical prote
161	6	2.2	50	2	H90982	H repeat-associate	234	6	2.2	142	2	T46542	hypothetical prote
162	6	2.2	61	1	B34913	adipokinetic hormo	235	6	2.2	143	2	T51530	40S RIBOSOMAL PROT
163	6	2.2	61	2	A27749	DNA-binding protei	236	6	2.2	143	2	S20937	photosystem I chai
164	6	2.2	62	2	A11756	hypothetical prote	237	6	2.2	143	2	A85333	hypothetical prote
165	6	2.2	64	2	T27749	DNA-binding protei	238	6	2.2	144	2	AF0663	conserved hypotet
166	6	2.2	67	2	C16372	hypothetical prote	239	6	2.2	144	2	C71149	Ig heavy chain V r
167	6	2.2	68	2	C83767	small acid-soluble	240	6	2.2	145	2	S11239	hemoglobin II - ar
168	6	2.2	68	2	H75273	conserved hypotet	241	6	2.2	146	1	HANK2	hemoglobin I - ar
169	6	2.2	69	2	A41903	recombinase homolo	242	6	2.2	146	1	GGNKTB	globin I - ark she
170	6	2.2	70	2	AC0287	conserved hypotet	243	6	2.2	147	2	S36144	hemoglobin I - ark
171	6	2.2	72	1	QQV29	hypothetical prote	244	6	2.2	147	2	AF1266	36S ribosomal prot
172	6	2.2	72	2	S15168	gas-vesicle protei	245	6	2.2	147	2	AF1081	50S ribosomal prot
173	6	2.2	72	2	E41903	recombinase Bin3 -	246	6	2.2	148	2	D87687	conserved hypotet
174	6	2.2	73	2	F83266	hypothetical prote	247	6	2.2	148	2	B95131	conserved domain p
175	6	2.2	75	2	T51499	hypothetical prote	248	6	2.2	148	2		

249	6	2.2	148	2	T31040	conserved hypot	322	6	2.2	185	2	S71512	hypothetical prote
250	6	2.2	149	2	F65169	o149 protein - Esc	323	6	2.2	186	2	G90541	gtp-binding protei
251	6	2.2	149	2	B91206	hypothetical prote	324	6	2.2	188	2	D70367	conserved hypot
252	6	2.2	149	2	D86052	hypothetical prote	325	6	2.2	188	2	AF2530	hypothetical prote
253	6	2.2	149	2	A13293	4-hydroxybenzoyl-C	326	6	2.2	189	1	IVHUF	interferon alpha-I
254	6	2.2	151	2	H70976	probable ribosoma	327	6	2.2	189	1	IVHUA5	interferon alpha-I
255	6	2.2	151	2	D86638	conserved hypot	328	6	2.2	189	1	IVHUA8	interferon alpha-I
256	6	2.2	152	2	H95153	conserved hypot	329	6	2.2	189	1	IVHUA9	interferon alpha-I
257	6	2.2	153	2	S60363	nucleoside-diphosp	330	6	2.2	189	1	IVHUA3	interferon alpha-7
258	6	2.2	153	2	S72982	ribosomal protein	331	6	2.2	189	1	IVHUI6	interferon alpha-I
259	6	2.2	153	2	A95928	probable amine oxi	332	6	2.2	189	1	IVHUA4	interferon alpha-4
260	6	2.2	154	2	P0C465	beta C protein - P	333	6	2.2	189	1	IVHUI8	interferon alpha-I
261	6	2.2	155	2	B71223	hypothetical prote	334	6	2.2	189	2	IS2347	interferon alpha-M
262	6	2.2	155	2	C75191	transcriptlon regu	335	6	2.2	189	2	IS4464	interferon-alpha-F
263	6	2.2	156	1	A49342	acetyl-CoA carboxy	336	6	2.2	189	2	IS1302	interferon-alpha-J
264	6	2.2	156	1	A64990	diheme cytochrome	337	6	2.2	189	2	IS1970	interferon precurs
265	6	2.2	156	2	D91015	cytochrome c-type	338	6	2.2	190	2	A99308	probable phosphoe
266	6	2.2	156	2	F85859	conserved hypot	339	6	2.2	190	2	T25201	hypothetical prote
267	6	2.2	156	2	B90508	hypothetical prote	340	6	2.2	192	2	S43563	RoH10.5 protein -
268	6	2.2	156	4	S59288	transcriptlon elon	341	6	2.2	192	2	S43611	competence transcr
269	6	2.2	157	2	AC0998	hypothetical prote	342	6	2.2	192	2	G83096	hypothetical prote
270	6	2.2	157	2	T23781	hypothetical prote	343	6	2.2	193	2	F70775	conserved hypot
271	6	2.2	158	2	D82782	protein K07C11.5 l	344	6	2.2	194	2	B75613	hypothetical prote
272	6	2.2	158	2	F89123	hypothetical prote	345	6	2.2	194	2	C75420	hypothetical prote
273	6	2.2	158	2	H83350	hypothetical prote	346	6	2.2	195	2	G75635	hypothetical prote
274	6	2.2	158	2	H85620	hypothetical prote	347	6	2.2	196	2	AR3096	phenylacrylic acid
275	6	2.2	159	2	AD0349	conserved hypot	348	6	2.2	196	2	D8190	decarboxylase (imp
276	6	2.2	159	2	AE1241	B. subtilis YqzC p	349	6	2.2	197	2	H75188	hypothetical prote
277	6	2.2	160	2	S65978	spore coat protein	350	6	2.2	198	1	QOECX3	putative methylase
278	6	2.2	161	1	F64447	conserved hypot	351	6	2.2	198	2	B91168	hypothetical prote
279	6	2.2	162	2	C25843	interferon alpha-B	352	6	2.2	198	2	B86014	hypothetical prote
280	6	2.2	163	2	S61520	hemoglobin alpha-B	353	6	2.2	198	2	AF0991	conserved hypot
281	6	2.2	165	2	S35287	gaf protein - phag	354	6	2.2	198	2	D72538	prphable [acyl-car
282	6	2.2	165	2	H64038	hypothetical prote	355	6	2.2	199	2	A69839	adenylylsulfate Ki
283	6	2.2	167	2	E25843	interferon alpha-F	356	6	2.2	199	2	I64125	nucleotide-binding
284	6	2.2	167	2	F25843	interferon alpha-J	357	6	2.2	199	2	D75442	probable peptidase
285	6	2.2	167	2	S49628	FUN81 protein - ye	358	6	2.2	200	1	RMMD4	ribosomal protein
286	6	2.2	168	2	AG3296	transcription regu	359	6	2.2	200	2	B95349	Probable decarboxy
287	6	2.2	168	2	B85740	hypothetical prote	360	6	2.2	200	2	F64068	recombination prot
288	6	2.2	168	2	E90878	probable filament	361	6	2.2	201	2	G69340	cobalamin biosynth
289	6	2.2	168	2	T17204	hypothetical prote	362	6	2.2	202	2	S68609	recombinase Sin -
290	6	2.2	168	2	G95023	acetyltransferase,	363	6	2.2	202	2	D56962	probable metal-dep
291	6	2.2	168	2	H97894	hypothetical prote	364	6	2.2	203	2	F90532	conserved hypot
292	6	2.2	168	2	E84362	GTP cyclohydrolase	365	6	2.2	203	2	T19760	hypothetical prote
293	6	2.2	170	2	A65136	transcription elon	366	6	2.2	204	2	G64061	ruvA protein - Hae
294	6	2.2	170	2	H91159	transcription elon	367	6	2.2	204	2	AE3302	multiple antibioti
295	6	2.2	170	2	G86005	transcription elon	368	6	2.2	205	2	E97019	competence protein
296	6	2.2	170	2	B95169	conserved domain p	369	6	2.2	206	2	AC3478	protein yb1s precu
297	6	2.2	170	2	D75400	conserved hypot	370	6	2.2	207	2	A10005	guanylate kinase (
298	6	2.2	170	2	C98035	hypothetical prote	371	6	2.2	207	2	T31959	hypothetical prote
299	6	2.2	171	2	S77242	hypothetical prote	372	6	2.2	207	2	T51567	hypothetical prote
300	6	2.2	173	2	T24942	hypothetical prote	373	6	2.2	208	1	RSYMAC	ribosomal protein
301	6	2.2	173	2	C86503	acetyltransferase (l	374	6	2.2	208	2	A38202	GMP-binding protei
302	6	2.2	176	2	I56314	interferon-alpha-	375	6	2.2	208	2	T43385	60S ribosomal prot
303	6	2.2	176	2	A72237	hypothetical prote	376	6	2.2	208	2	A38594	tropoin I - fruit
304	6	2.2	176	2	A83050	outer membrane lip	377	6	2.2	208	2	AA0547	tropoin I - fruit
305	6	2.2	178	2	D71668	gripe protein (GRP5	378	6	2.2	209	2	F83149	thiamin-phosphate
306	6	2.2	178	2	A71939	hypothetical prote	379	6	2.2	209	2	D88577	protein RoH10.5 (
307	6	2.2	178	2	S64648	hypothetical prote	380	6	2.2	210	1	D64238	hypothetical prote
308	6	2.2	178	2	S67379	hypothetical prote	381	6	2.2	210	2	F71308	probable 2-dehydro
309	6	2.2	179	2	T49757	hypothetical prote	382	6	2.2	210	2	G76316	hypothetical prote
310	6	2.2	179	2	T51570	hypothetical prote	383	6	2.2	210	2	F83393	probable transcrip
311	6	2.2	180	2	T41322	probable succinate	384	6	2.2	210	2	E75349	hypothetical prote
312	6	2.2	180	2	S58767	probable type 1 fi	385	6	2.2	212	2	T36864	probable cyclohexa
313	6	2.2	180	2	S58767	stretptochiricin ace	386	6	2.2	212	2	C55544	flbE protein - Cau
314	6	2.2	180	2	H84154	hypothetical prote	387	6	2.2	213	2	B89820	hypothetical prote
315	6	2.2	180	2	D90569	hypothetical prote	388	6	2.2	213	2	P98019	hypothetical prote
316	6	2.2	181	2	I56313	interferon alpha 2	389	6	2.2	214	2	B95898	probable sensory t
317	6	2.2	181	2	E82077	ampd protein VC242	390	6	2.2	215	2	S61337	proteasome endopep
318	6	2.2	181	2	T35851	hypothetical prote	391	6	2.2	216	2	G27362	hypothetical prote
319	6	2.2	182	2	G70698	peptidylprolyl iso	392	6	2.2	217	2	B97948	hypothetical prote
320	6	2.2	185	2	S67036	SP2 protein - yea	393	6	2.2	218	2	H64437	TRK system potassi
321	6	2.2	185	2	C82662	hypothetical prote	394	6	2.2	218	2	B64365	hypothetical prote

395	2.2	219	2	D82603	hypotheical prote	468	2.2	254	2	B81293	hypotheical prote
396	2.2	220	2	H69257	hypotheical prote	469	2.2	255	2	B81929	probable imidazole
397	2.2	220	2	T35744	hypotheical prote	470	2.2	255	2	H81176	hlf protein NMB06
398	2.2	221	2	A39491	conserved hypotet	471	2.2	255	2	S74930	hypotheical prote
399	2.2	222	2	A56276	H+-transporting ox	472	2.2	255	2	D64833	probable ABC-type
400	2.2	225	2	T11471	hypotheical prote	473	2.2	255	2	H85619	probable ABC-type
401	2.2	226	2	S10496	hypotheical prote	474	2.2	255	2	H90755	probable ABC-type
402	2.2	226	2	AG1883	probable fibrillar	475	2.2	256	2	G75404	outer membrane pro
403	2.2	227	2	AB1076	hypotheical prote	476	2.2	256	2	H75404	hypotheical prote
404	2.2	228	2	T47999	hypotheical prote	477	2.2	257	2	T57955	probable Spou-fam1
405	2.2	228	2	E69517	hypotheical prote	478	2.2	257	2	AB0353	haptoglobin Hp - c
406	2.2	228	2	E770934	probable lpgn prot	479	2.2	258	2	T16945	haptoglobin Hsp - c
407	2.2	229	2	E71010	hypotheical prote	480	2.2	258	2	G02959	haptoglobin - fms
408	2.2	229	2	T15359	hypotheical prote	481	2.2	258	2	G24812	similar to Mtn3 pr
409	2.2	230	2	A61607	probable hemolysin	482	2.2	259	1	B64251	probable methyltra
410	2.2	231	2	G72463	hypotheical prote	483	2.2	259	1	P06687	cyclin - anthracro
411	2.2	231	2	J01601	replication protei	484	2.2	259	2	CG3153	conserved hypotet
412	2.2	231	2	C70000	ABC transporter (A	485	2.2	259	2	CG4437	hypotheical prote
413	2.2	232	2	DXCH	ovalbumin-related	486	2.2	259	2	AB3238	hypotheical prote
414	2.2	232	2	A95081	hypotheical prote	487	2.2	259	2	AB3238	oligopeptide trans
415	2.2	237	2	F82178	probable fibrillar	488	2.2	259	2	AG8302	hypotheical prote
416	2.2	237	2	C97948	hypotheical prote	489	2.2	259	2	AG2981	chlorocatechol 1,2
417	2.2	239	2	A12773	sugar fermentation	490	2.2	260	1	A27058	chlorocatechol 1,2
418	2.2	239	2	G97553	hypotheical prote	491	2.2	260	2	T44616	chlorocatechol 1,2
419	2.2	239	2	TS1565	hypotheical prote	492	2.2	260	2	T44666	hypotheical prote
420	2.2	239	2	G90366	hypotheical prote	493	2.2	260	2	B90026	probable enoyl-CoA
421	2.2	239	2	AH2836	O-methyltransferas	494	2.2	261	2	H83442	glutamine-binding
422	2.2	239	2	B97614	hypotheical prote	495	2.2	262	2	B42478	probable periplasm
423	2.2	240	2	G91045	hypotheical prote	496	2.2	262	2	B81326	probable undecapre
424	2.2	240	2	C85890	hypotheical prote	497	2.2	262	2	H87217	hypotheical prote
425	2.2	240	2	H65002	hypotheical prote	498	2.2	262	2	T14807	conserved hypotet
426	2.2	241	2	AP0758	precorrin-3B C17-m	499	2.2	264	2	H83351	hypotheical prote
427	2.2	242	2	S19267	anthranilate phosp	500	2.2	264	2	AE1362	probable terminase
428	2.2	242	2	T37107	hypotheical prote	501	2.2	264	2	AE1362	hypotheical prote
429	2.2	242	2	AB1091	autolysin (N-acety	502	2.2	265	2	G84049	molysiderum ABC tra
430	2.2	242	2	A11454	arginine-binding p	503	2.2	265	2	C64579	hypotheical prote
431	2.2	243	2	AG0607	probable RNA methy	504	2.2	265	2	G71934	molysiderum ABC tran
432	2.2	243	2	AH0824	capsula polysacch	505	2.2	267	2	T24397	hypotheical prote
433	2.2	243	2	UC5723	conserved membrane	506	2.2	267	2	A69867	conserved hypotet
434	2.2	243	2	C96909	hypotheical prote	507	2.2	267	2	A75217	hypotheical prote
435	2.2	244	2	T20784	hypotheical prote	508	2.2	267	2	D95089	peptidyl-prolyl ci
436	2.2	244	2	A90302	oxidoreductase 5'-phosp	509	2.2	267	2	G97956	dipeptide transpor
437	2.2	245	2	AP0271	phosphatidyl - cat	510	2.2	268	2	A71952	dipeptide transpor
438	2.2	245	2	S52096	probable ATP synth	511	2.2	268	2	F64557	TomB-dependent rec
439	2.2	245	2	P91053	probable ATP synth	512	2.2	268	2	C87579	ABC transporter su
440	2.2	245	2	B85888	conserved hypotet	513	2.2	268	2	AD3496	hypotheical prote
441	2.2	245	2	S66095	hypotheical prote	514	2.2	268	2	AE4501	hypotheical prote
442	2.2	245	2	C83584	major antigenic pe	515	2.2	268	2	AE4501	beta-lactamase reg
443	2.2	245	2	B81349	dipeptide transpor	516	2.2	269	1	H75503	cobalam synthase
444	2.2	245	2	A99163	hypotheical prote	517	2.2	269	2	H75576	probable aliphatic
445	2.2	245	2	AH3124	hypotheical prote	518	2.2	271	2	AD0441	lipoprotein-28 pre
446	2.2	246	1	C65030	hypotheical prote	519	2.2	272	1	LEPC28	probable dehydroge
447	2.2	246	2	B81305	probable membrane	520	2.2	272	2	E70879	lipoprotein-28 (im
448	2.2	246	2	H87427	hypotheical prote	521	2.2	272	2	CG1203	lipoprotein-28 (im
449	2.2	246	2	F97076	hypotheical prote	522	2.2	272	2	D66049	lipoprotein, attac
450	2.2	247	2	PC4260	activin type I rec	523	2.2	273	2	G97021	proteaseome endope
451	2.2	248	2	T26461	hypotheical prote	524	2.2	273	2	T51977	proteaseome endope
452	2.2	248	2	B95203	conserved hypotet	525	2.2	273	2	H6370	33.2K hypotet
453	2.2	248	2	B98070	conserved hypotet	526	2.2	274	2	T51979	proteaseome endope
454	2.2	249	1	R5B77	60S ribosomal prot	527	2.2	274	2	CG1914	hypotheical prote
455	2.2	249	2	AF3434	surfeit locus prot	528	2.2	274	2	H81187	conserved hypotet
456	2.2	249	2	A97154	uncharacterized pt	529	2.2	274	2	G83214	probable ATP-bindi
457	2.2	249	2	T21920	hypotheical prote	530	2.2	274	2	C66213	hypotheical prote
458	2.2	249	2	F71342	probable UDP-N-ace	531	2.2	274	2	AB2509	probable phosphati
459	2.2	251	2	S52989	superoxide dismuta	532	2.2	275	2	T42747	hypotheical prote
460	2.2	251	2	A44506	lactose operon rep	533	2.2	275	2	S76778	hypotheical prote
461	2.2	251	2	G90015	conserved hypotet	534	2.2	275	2	B83626	transcription regu
462	2.2	251	2	F83101	hypotheical prote	535	2.2	276	2	B83983	hypotheical prote
463	2.2	251	2	A84367	hypotheical prote	536	2.2	276	2	H83568	probable permease
464	2.2	252	2	C89825	dehydrogenase homo	537	2.2	276	2	H75169	abc transporter PA
465	2.2	253	2	AG1346	ABC transporter, A	538	2.2	277	2	AC3305	23S ribosomal RNA
466	2.2	253	2	E69497	ABC transporter, A	539	2.2	277	2	C93358	probable oxidoredu
467	2.2	254	2	A72801	gp12 protein - Myc	540	2.2	277	2	G71849	amino acid ABC tra

541	6	2.2	277	2	D64666	614	6	2.2	308	2	AF1581	B. subtilis ribonu
542	6	2.2	278	2	B82388	615	6	2.2	309	1	XYECM	homoserine O-succi
543	6	2.2	279	2	S27149	616	6	2.2	309	1	XYECM	carbamate kinase (
544	6	2.2	279	2	A87108	617	6	2.2	309	2	C01245	homoserine transu
545	6	2.2	280	2	PH0269	618	6	2.2	309	2	A86093	homoserine transu
546	6	2.2	280	2	A89612	619	6	2.2	309	2	AC1011	homoserine O-succi
547	6	2.2	280	2	A83796	620	6	2.2	309	2	AD0502	homoserine kinase
548	6	2.2	280	2	D75077	621	6	2.2	309	2	A12809	hypothetical prote
549	6	2.2	282	2	H84653	622	6	2.2	309	2	A90439	ABC transporter, A
550	6	2.2	283	2	JC5661	623	6	2.2	309	2	T52453	ATP-dependent Clp
551	6	2.2	284	2	B95017	624	6	2.2	310	2	CE9218	transcription init
552	6	2.2	284	2	C97890	625	6	2.2	310	2	T52041	probable ATP-depen
553	6	2.2	285	1	C70873	626	6	2.2	311	2	D64209	hydroxymethylgluta
554	6	2.2	285	2	A95851	627	6	2.2	312	2	AC2904	oxidoreductase Atu
555	6	2.2	286	2	JC6007	628	6	2.2	312	2	E97679	probable oxidoredu
556	6	2.2	286	2	AB1234	629	6	2.2	313	2	AD0920	porphobilinogen de
557	6	2.2	286	2	AB1596	630	6	2.2	313	2	CE9788	hypothetical prote
558	6	2.2	286	2	G84898	631	6	2.2	313	2	A69496	trp domain-contain
559	6	2.2	286	2	E70124	632	6	2.2	314	1	DEPXM	malate dehydrogena
560	6	2.2	287	2	C81023	633	6	2.2	314	1	A41671	iron transport pro
561	6	2.2	287	2	G81968	634	6	2.2	315	1	YXCKTA	thymidylate syntha
562	6	2.2	287	2	AD0006	635	6	2.2	315	2	H96598	protein F20N2.8 [i
563	6	2.2	289	2	S12619	636	6	2.2	316	2	JC4701	cadmium, zinc, cob
564	6	2.2	290	2	AB1862	637	6	2.2	316	2	D71375	probable ABC trans
565	6	2.2	291	2	E84408	638	6	2.2	316	2	A41626	protein farnesyltr
566	6	2.2	291	2	F69959	639	6	2.2	317	2	E95239	conserved hypotnet
567	6	2.2	292	2	T16085	640	6	2.2	317	2	E98103	conserved exonuclea
568	6	2.2	293	2	H65182	641	6	2.2	317	2	E86419	probable lacyl-car
569	6	2.2	293	2	B91219	642	6	2.2	318	2	T7548	probable kinase
570	6	2.2	293	2	D86065	643	6	2.2	318	2	G82086	unknown protein Fl
571	6	2.2	293	2	D81204	644	6	2.2	318	2	D96774	homeodomain transc
572	6	2.2	293	2	A81780	645	6	2.2	318	2	T00402	probable ABC trans
573	6	2.2	293	2	D95919	646	6	2.2	320	2	G95268	probable oxidoredu
574	6	2.2	293	2	T31146	647	6	2.2	321	2	T38413	oligopeptide ABC t
575	6	2.2	294	1	G65106	648	6	2.2	321	2	G96921	ethyrocyste band 7
576	6	2.2	294	2	G72293	649	6	2.2	322	2	S74505	3-dehydroquinatase
577	6	2.2	294	2	T10690	650	6	2.2	322	2	A83857	triose-phosphate i
578	6	2.2	294	2	D91134	651	6	2.2	322	2	S52032	hypothetical prote
579	6	2.2	294	2	AB0424	652	6	2.2	322	2	G64388	hypothetical prote
580	6	2.2	294	2	G85979	653	6	2.2	322	2	G72643	probable transkeo
581	6	2.2	294	2	AC0901	654	6	2.2	323	2	A64054	selenophosphate sy
582	6	2.2	296	1	JQ2155	655	6	2.2	323	2	D83708	phosphonates trans
583	6	2.2	296	1	T51326	656	6	2.2	324	2	AC0088	putative flagellar
584	6	2.2	296	2	T46619	657	6	2.2	325	2	F82068	2-hydroxyacid dehy
585	6	2.2	297	2	T22690	658	6	2.2	325	2	T07001	cysteine synthase
586	6	2.2	297	2	T34165	659	6	2.2	325	2	C72289	oligopeptide ABC t
587	6	2.2	298	1	C42802	660	6	2.2	326	2	T25728	hypothetical prote
588	6	2.2	298	1	D69023	661	6	2.2	326	2	B69412	transcription init
589	6	2.2	298	2	B83553	662	6	2.2	327	2	B67168	probable membrane
590	6	2.2	298	2	T24029	663	6	2.2	328	2	I64164	hypothetical prote
591	6	2.2	298	2	D95987	664	6	2.2	329	1	HPDG	haptoglobin precu
592	6	2.2	299	2	A95878	665	6	2.2	329	2	AC0942	periplasmic sulpha
593	6	2.2	299	2	T37756	666	6	2.2	330	2	S67139	probable membrane
594	6	2.2	300	2	B64508	667	6	2.2	330	2	C83995	branched-chain alp
595	6	2.2	300	2	H87390	668	6	2.2	330	2	C69593	3-methyl-2-oxobuta
596	6	2.2	300	2	F69719	669	6	2.2	332	2	C97111	glycerol 3-phospha
597	6	2.2	300	2	AC3175	670	6	2.2	332	2	A46408	abscisic acid-indu
598	6	2.2	301	2	AB3195	671	6	2.2	332	2	A36663	neutrlus protein -
599	6	2.2	301	2	S51132	672	6	2.2	333	2	H69679	involved in fatty
600	6	2.2	302	2	A89898	673	6	2.2	333	2	AE0266	oligopeptide trans
601	6	2.2	302	2	A40369	674	6	2.2	333	2	PQ0377	hemagglutinin - me
602	6	2.2	303	1	GEPCO	675	6	2.2	334	2	OREBOF	oligopeptide trans
603	6	2.2	303	2	H71693	676	6	2.2	334	2	H71524	probable aspartate
604	6	2.2	304	2	E97243	677	6	2.2	334	2	B64872	oligopeptide trans
605	6	2.2	305	2	B97762	678	6	2.2	334	2	B85705	hypothetical prote
606	6	2.2	305	2	A70482	679	6	2.2	334	2	C90847	hypothetical prote
607	6	2.2	305	2	F86456	680	6	2.2	334	2	AC0651	oligopeptide trans
608	6	2.2	306	2	E83169	681	6	2.2	335	2	F71211	probable oligopept
609	6	2.2	306	2	T46757	682	6	2.2	335	2	F75026	oligopeptide trans
610	6	2.2	306	2	AG2104	683	6	2.2	335	2	AG0183	galactose-binding
611	6	2.2	306	2	S27496	684	6	2.2	335	2	G81280	holiday junction
612	6	2.2	308	2	S24802	685	6	2.2	335	2	A41845	off A - Treponema
613	6	2.2	308	2	AD1228	686	6	2.2	335	2	A72289	oligopeptide ABC t

687	6	2.2	335	2	AD3492	760	6	2.2	354	2	T52401	branded-chain ami
688	6	2.2	336	1	A64500	761	6	2.2	354	2	C90882	probable transcript
689	6	2.2	336	1	F82242	762	6	2.2	354	2	F85736	probable transcript
690	6	2.2	336	2	F71942	763	6	2.2	354	2	A64894	probable regulator
691	6	2.2	336	2	C64652	764	6	2.2	354	2	A12889	lvsyl-tRNA synthet
692	6	2.2	336	2	E84936	765	6	2.2	354	2	F97665	NodI membrane tran
693	6	2.2	336	2	T00832	766	6	2.2	355	2	H85320	heat inducible tra
694	6	2.2	336	2	T21565	767	6	2.2	355	2	E97928	NADH dehydrogenase
695	6	2.2	336	2	H69437	768	6	2.2	356	2	D87450	coenzyme PQQ synth
696	6	2.2	337	2	C70990	769	6	2.2	356	2	D84272	probable sulfite o
697	6	2.2	338	1	DEPTMM	770	6	2.2	357	2	T29856	probable aspartate
698	6	2.2	338	1	DEMSM	771	6	2.2	357	2	S76088	hypothetical prote
699	6	2.2	338	2	T03432	772	6	2.2	357	2	T01697	probable triacylg
700	6	2.2	338	2	F83965	773	6	2.2	357	2	T17919	hypothetical prote
701	6	2.2	338	2	A83835	774	6	2.2	357	2	T17919	protein co-factor
702	6	2.2	339	2	A41677	775	6	2.2	358	2	I46532	decorin precursor
703	6	2.2	339	2	AC3245	776	6	2.2	359	1	NBHUC8	decorin precursor
704	6	2.2	339	2	T20442	777	6	2.2	359	2	T21705	hypothetical prote
705	6	2.2	339	2	T27715	778	6	2.2	359	2	AD2743	conserved hypochet
706	6	2.2	340	2	F70616	779	6	2.2	359	2	C97524	hypothetical prote
707	6	2.2	340	2	JH0363	780	6	2.2	360	2	S06280	decorin precursor
708	6	2.2	340	2	B41969	781	6	2.2	360	2	I47020	decorin - fibrill
709	6	2.2	340	2	E69544	782	6	2.2	360	2	H84401	hypothetical prote
710	6	2.2	340	2	D98346	783	6	2.2	361	2	T39784	hypothetical prote
711	6	2.2	341	2	G71313	784	6	2.2	361	2	AC2936	hypothetical prote
712	6	2.2	341	2	F87820	785	6	2.2	362	2	G95029	glycerol dehydroge
713	6	2.2	342	2	A30189	786	6	2.2	362	2	B97901	glycerol dehydroge
714	6	2.2	342	2	S75086	787	6	2.2	362	2	C82572	phosphoserine amin
715	6	2.2	342	2	S57510	788	6	2.2	363	2	G95937	probable mureinpep
716	6	2.2	342	2	C71372	789	6	2.2	364	2	G84367	hypothetical prote
717	6	2.2	342	2	A83867	790	6	2.2	365	2	D89980	conserved hypochet
718	6	2.2	343	2	D91032	791	6	2.2	367	2	T17481	prephenate dehydro
719	6	2.2	343	2	E85976	792	6	2.2	367	2	G83309	conserved hypochet
720	6	2.2	344	2	S42648	793	6	2.2	368	2	JH0348	VSG expression sit
721	6	2.2	344	2	AB2306	794	6	2.2	368	2	G86416	hypothetical prote
722	6	2.2	344	2	S76268	795	6	2.2	369	1	A33559	larnesyltransfarn
723	6	2.2	344	2	T05524	796	6	2.2	369	2	S53722	hypothetical prote
724	6	2.2	344	2	H95059	797	6	2.2	369	2	S76784	hypothetical prote
725	6	2.2	344	2	AG1005	798	6	2.2	369	2	D87476	hypothetical prote
726	6	2.2	345	2	F95216	799	6	2.2	369	2	S75038	hypothetical prote
727	6	2.2	345	2	D98080	800	6	2.2	369	2	T16506	glycerol dehydroge
728	6	2.2	345	2	I36941	801	6	2.2	370	1	U01474	glycerol dehydroge
729	6	2.2	345	2	S09858	802	6	2.2	370	2	A86737	(R,R)-butanediol d
730	6	2.2	345	2	A10830	803	6	2.2	370	2	T11849	gibberellin 20-oxi
731	6	2.2	345	2	C95873	804	6	2.2	370	2	B72018	8-amino-7-oxononan
732	6	2.2	345	2	AH3131	805	6	2.2	370	2	A86506	oxononanoate synth
733	6	2.2	345	2	AD3635	806	6	2.2	370	2	C72290	branched chain ami
734	6	2.2	345	2	C90416	807	6	2.2	371	2	D75402	probable N-acetyl
735	6	2.2	345	2	G83633	808	6	2.2	372	2	G82181	galactose-1-phosph
736	6	2.2	346	2	T19834	809	6	2.2	372	2	A52061	hypothetical prote
737	6	2.2	346	2	I36942	810	6	2.2	373	2	H84187	sarcosine oxidase
738	6	2.2	346	2	C81446	811	6	2.2	373	2	B84595	hypothetical prote
739	6	2.2	346	2	C98156	812	6	2.2	373	2	A69773	hypothetical prote
740	6	2.2	346	2	T19676	813	6	2.2	375	2	T39364	probable galactosy
741	6	2.2	347	1	HPHUT	814	6	2.2	375	2	C95106	carboxymorspermid
742	6	2.2	347	1	HPMT	815	6	2.2	375	2	A13383	hypothetical prote
743	6	2.2	347	1	HPRT	816	6	2.2	375	2	T32251	hypothetical prote
744	6	2.2	347	1	F90937	817	6	2.2	375	2	E87974	carboxymorspermid
745	6	2.2	347	2	B85786	818	6	2.2	376	2	T02770	glutathione-disulf
746	6	2.2	347	2	JM0033	819	6	2.2	376	2	E35879	probable sugar ABC
747	6	2.2	347	2	AH0710	820	6	2.2	377	2	F97250	proline/glycine be
748	6	2.2	347	2	T38154	821	6	2.2	377	2	A38662	tRNA-splicing eno
749	6	2.2	347	2	B96008	822	6	2.2	377	2	D68881	pyridine nucleotid
750	6	2.2	347	2	D97588	823	6	2.2	378	1	OKGAK1	protein kinase (EC
751	6	2.2	348	2	HPHUR	824	6	2.2	378	2	C70119	hypothetical prote
752	6	2.2	349	2	I36944	825	6	2.2	380	2	T00683	hypothetical prote
753	6	2.2	350	2	A42009	826	6	2.2	380	2	T48052	hypothetical prote
754	6	2.2	350	2	F75448	827	6	2.2	381	2	S74777	hypothetical prote
755	6	2.2	352	1	CHECFX	828	6	2.2	381	2	JC4378	metalloprotease
756	6	2.2	352	1	G91167	829	6	2.2	381	2	T38655	hypothetical 37.5
757	6	2.2	352	2	G86013	830	6	2.2	383	2	S11433	ovalbumin - Japane
758	6	2.2	353	2	T03687	831	6	2.2	385	2	G87340	anthranilate
759	6	2.2	353	2	A10220	832	6	2.2	385	2	A44102	di-N-acetylchitobi

833	6	2.2	385	1	OACH	906	6	2.2	417	2	A13382	lysine-tRNA ligase
834	6	2.2	387	2	C92336	907	6	2.2	417	2	S74940	hypothetical prote
835	6	2.2	387	2	A83072	908	6	2.2	417	2	S56454	hypothetical ABC t
836	6	2.2	387	2	ORF MSV41 leucine	909	6	2.2	418	2	C83341	serine hydroxymeth
837	6	2.2	388	1	DYCH	910	6	2.2	419	2	AF0104	maltoxin [import
838	6	2.2	388	2	B48899	911	6	2.2	419	2	E95157	Atz/712 family pro
839	6	2.2	388	2	T26977	912	6	2.2	420	1	AUBIRS	argininosuccinate
840	6	2.2	388	2	AC0138	913	6	2.2	420	2	T34750	serine hydroxymeth
841	6	2.2	389	2	TS0671	914	6	2.2	420	2	A96535	unknown protein, 1
842	6	2.2	390	2	AF0426	915	6	2.2	420	2	E96614	hypothetical prote
843	6	2.2	390	2	T04036	916	6	2.2	421	1	D64100	glycine hydroxymet
844	6	2.2	390	2	AC0112	917	6	2.2	421	2	JY0057	tolA protein -Bac
845	6	2.2	392	2	D53240	918	6	2.2	422	2	G84073	C4-dicarboxylate t
846	6	2.2	392	2	DB6823	919	6	2.2	423	2	S86808	hypothetical prote
847	6	2.2	393	2	C83049	920	6	2.2	424	2	F95894	probable nitrate t
848	6	2.2	393	2	A71154	921	6	2.2	424	2	A95159	psr protein [import
849	6	2.2	393	2	D97275	922	6	2.2	425	2	S53004	conserved hypothet
850	6	2.2	394	2	G84094	923	6	2.2	425	2	F71871	mitosis-specific c
851	6	2.2	394	2	F90725	924	6	2.2	426	2	C70896	topoisomerase I -
852	6	2.2	394	2	G85576	925	6	2.2	426	2	B87463	glycine hydroxymet
853	6	2.2	395	2	B72381	926	6	2.2	428	2	D70583	hypothetical prote
854	6	2.2	395	2	T38685	927	6	2.2	428	2	B89902	probable transmem
855	6	2.2	396	2	D81408	928	6	2.2	429	2	T45520	transposase [import
856	6	2.2	396	2	A39039	929	6	2.2	429	2	S29044	endothucanase A pr
857	6	2.2	396	2	T04100	930	6	2.2	429	2	AD2873	hypothetical prote
858	6	2.2	396	2	A84556	931	6	2.2	430	2	G70352	N-ethylameline ch
859	6	2.2	396	2	AE2137	932	6	2.2	430	2	S55325	endo-beta-1,6-gluc
860	6	2.2	397	2	C53240	933	6	2.2	432	2	D55333	hypothetical prote
861	6	2.2	397	2	C39099	934	6	2.2	433	2	E96649	hypothetical prote
862	6	2.2	397	2	A64014	935	6	2.2	435	2	A81416	probable integrat
863	6	2.2	397	2	T19182	936	6	2.2	435	2	BE4079	putrescine/cornubi
864	6	2.2	397	2	A84064	937	6	2.2	435	2	T01805	hypothetical prote
865	6	2.2	398	2	E86655	938	6	2.2	436	2	E64340	hypothetical prote
866	6	2.2	398	2	S29819	939	6	2.2	436	2	T39883	hypothetical prote
867	6	2.2	398	2	T34847	940	6	2.2	437	1	D64457	translaton initia
868	6	2.2	398	2	B53340	941	6	2.2	437	2	I39578	nccc protein - Alc
869	6	2.2	398	2	B39099	942	6	2.2	437	2	H96933	aspartate kinase [
870	6	2.2	398	2	B97341	943	6	2.2	439	2	D70958	hypothetical prote
871	6	2.2	399	2	C71728	944	6	2.2	439	2	B86302	hypothetical prote
872	6	2.2	399	2	T28359	945	6	2.2	439	2	H70943	hypothetical prote
873	6	2.2	400	2	A47094	946	6	2.2	440	2	I54442	peroxisome prolife
874	6	2.2	400	2	D86575	947	6	2.2	440	2	UC4530	probable MFS trans
875	6	2.2	402	2	T72949	948	6	2.2	440	2	AB1015	probable DNA-damag
876	6	2.2	402	2	TI4438	949	6	2.2	441	2	F68185	hypothetical prote
877	6	2.2	402	2	B98175	950	6	2.2	443	1	A46248	dihydroorotate oxi
878	6	2.2	403	2	C70832	951	6	2.2	443	2	S43899	methyl coenzyme M
879	6	2.2	403	2	H96026	952	6	2.2	445	2	S83743	N-ethylameline ch
880	6	2.2	405	2	S64052	953	6	2.2	445	2	T28015	hypothetical prote
881	6	2.2	405	2	C83204	954	6	2.2	445	2	I46225	probable chloropla
882	6	2.2	405	2	E81946	955	6	2.2	445	2	E97649	hypothetical prote
883	6	2.2	405	2	F82229	956	6	2.2	445	2	T35005	probable integrat
884	6	2.2	406	1	HPH2	957	6	2.2	447	2	S32227	glutamate dehydrog
885	6	2.2	406	2	C81415	958	6	2.2	447	2	T02547	hypothetical prote
886	6	2.2	407	2	H81160	959	6	2.2	447	2	G96569	hypothetical prote
887	6	2.2	408	2	C91062	960	6	2.2	447	2	D85646	hypothetical prote
888	6	2.2	408	2	G85906	961	6	2.2	447	2	F90786	hypothetical prote
889	6	2.2	408	2	G65038	962	6	2.2	447	2	T18264	cellulosome anchor
890	6	2.2	409	1	F68614	963	6	2.2	448	2	G70009	organic acid trans
891	6	2.2	410	2	C95971	964	6	2.2	448	2	H97249	protein containing
892	6	2.2	410	2	H95888	965	6	2.2	451	2	B96495	phosphoglucomutase
893	6	2.2	410	2	AH1484	966	6	2.2	451	2	H82044	C4-dicarboxylate t
894	6	2.2	410	2	AE3092	967	6	2.2	451	2	S57588	hypothetical prote
895	6	2.2	410	2	E98194	968	6	2.2	451	2	T30732	phosphomevalonate
896	6	2.2	412	1	G30315	969	6	2.2	451	2	C86766	lambda-type inclusion b
897	6	2.2	412	2	E70108	970	6	2.2	451	2	B96002	hypothetical prote
898	6	2.2	412	2	S62538	971	6	2.2	451	2	B91280	conserved hypothet
899	6	2.2	412	2	T41552	972	6	2.2	457	2	B86121	probable ligase [i
900	6	2.2	412	2	S66925	973	6	2.2	457	2	S56459	UDP-N-acetylmurama
901	6	2.2	414	2	A10156	974	6	2.2	457	2	AC1056	muirein peptide lig
902	6	2.2	414	2	T34815	975	6	2.2	457	2		
903	6	2.2	414	2	AG0075	976	6	2.2	457	2		
904	6	2.2	416	2	A82892	977	6	2.2	457	2		
905	6	2.2	417	2	G82968	978	6	2.2	457	2		

C:Superfamily: lipoprotein-28

Query Match 4.0%; Score 11; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 FVEDKDSPPVN 234
Db 219 FVEDKDSPPVN 229

RESULT 5

G85504
C:Species: lipoprotein yaeC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: G85504
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, L.; Grobbeck, E.U.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: G85504
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-271 <STO>
A/Cross-references: GB:AE005174; NID:912512928; PIDN:AA654499.1; GSPDB:GN00145; UWGP:Z02
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: yaeC
C:Superfamily: lipoprotein-28

Query Match 4.0%; Score 11; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 FVEDKDSPPVN 234
Db 219 FVEDKDSPPVN 229

RESULT 6

A10532
C:Species: lipoprotein precursor STY0272 [imported] - Salmonella enterica subsp. enterica
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: A10532
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A/Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21594947; PMID:11677608
A/Accession: A10532
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-271 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD08705.1; PID:G16501528; GSPDB:GN00176
C/Genetics:
A/Gene: STY0272
C:Superfamily: lipoprotein-28

Query Match 4.0%; Score 11; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 FVEDKDSPPVN 234
Db 219 FVEDKDSPPVN 229

RESULT 7

A48107
A/Title: RNA polymerase (EC 2.7.7.6) I chain A12.2 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein J1747; protein YJR063w
C/Species: Saccharomyces cerevisiae
C/Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 22-Jun-2003
C/Accession: A48107; S47937; S57082; S71685
R:Nogi, Y.; Yano, R.; Dodd, J.; Carles, C.; Nomura, M.
Mol. Cell. Biol. 13, 114-122, 1993
A/Title: Gene RNA4 in Saccharomyces cerevisiae encodes the A12.2 subunit of RNA polymerase
A/Reference number: A48107; MUID:93109294; PMID:8417319
A/Accession: A48107
A/Molecule type: DNA
A/Residues: 1-125 <NOG>

A/Cross-references: EMBL:L00708; NID:9172461; PIDN:AAA3992.1; PID:G172462
A/Note: Sequence extracted from NCBI backbone (NCBIN:121124, NCBI:121126)
A/Accession: B48107
A/Molecule type: DNA
A/Residues: 40-46; 48-61; 104-113 <NO2>
R:Baker, R.T.; Varshavsky, A.
submitted to the EMBL Data Library, September 1994
A/Description: N-terminal amide: a new enzyme and component of a targeting complex in t
A/Reference number: S47937
A/Accession: S47937
A/Molecule type: DNA
A/Residues: 1-125 <BAK>

A/Cross-references: EMBL:U35664; NID:9531230; PIDN:AA859319.1; PID:G531231
R:Huang, M.E.; Chuat, J.C.; Galibert, F.
submitted to the Protein Sequence Database, September 1995
A/Reference number: S57052
A/Accession: S57052
A/Molecule type: DNA
A/Residues: 1-125 <MAN>

A/Cross-references: EMBL:U49563; NID:91015736; PIDN:CAA89591.1; PID:G1015737; MIPS:YJR06;
R:Huang, M.E.; Manns, V.; Chuat, J.C.; Galibert, F.
Yeast 12, 865-875, 1996
A/Title: Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open reading frames
A/Reference number: S71676; MUID:96437976; PMID:8840504
A/Accession: S71685
A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
A/Residues: 1-125 <HUA>
A/Cross-references: EMBL:U47993; NID:91019675; PIDN:AA839289.1; PID:G1019665
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C/Genetics:
A/Gene: SGD:RPA12; RNA4
A/Cross-references: SGD:S0003824; MIPS:YJR063w
A/Map position: 108
C:Superfamily: DNA-directed RNA polymerase chain A12.2/B12.6/ C11/M/14.5K
C/Keywords: nucleotidyltransferase; nucleus; transcription

Query Match 3.3%; Score 9; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 132 NELKDGATI 140
Db 74 NELKDGATI 82

RESULT 8

A1110
C:Species: lipoprotein lmo0285 [imported] - Listeria monocytogenes (strain EGD-e)
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: A1110
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baguer, F.; Berche, P.; Blocker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.;
Science 294, 849-852, 2001

A:Reference number: AB1077; MWID:21537279; PMID:11679669
A:Accession: AF110
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-273 <GUA>
A:Cross-references: GB:NC_00310; PIDN:CAD00812.1; PID:916406550; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0285
C:Superfamily: lipo-protein-28

Query Match
Best Local Similarity 3.3%; Score 9; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 IAVPNDPSN 148
Db 136 IAVPNDPSN 144

RESULT 9
A:Accession: A82266
A:Protein: YaeC VC0905 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82266
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chaudson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F. 1, R.R.; Mekalanos, J.D.; Venter, J.C.; Fraser, C.M.
A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MWID:20406833; PMID:10952301
A:Accession: A82266
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-275 <HEI>
A:Cross-references: GB:AE004174; GB:AE003852; NID:99655355; PIDN:AAF94067.1; GSPDB:GN001
A:Experimental source: serogroup O1, strain N16961, biotype El Tor
C:Genetics:
A:Gene: VC0905
A:Map position: 1
C:Superfamily: lipo-protein-28

Query Match
Best Local Similarity 3.3%; Score 9; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 IVGNFVYP 120
Db 111 IVGNFVYP 119

RESULT 10
A:Accession: A13551
ABC transporter substrate binding protein BME10338 [Imported] - *Brucella melitensis* (str. C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: A13551
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess, Proc. Natl. Acad. Sci. U.S.A. 95, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: A13551
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-278 <KUR>
A:Cross-references: GB:AE00918; PIDN:AAU53580.1; PID:917984491; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME10338
A:Map position: 11
C:Superfamily: lipo-protein-28

Query Match
Best Local Similarity 3.3%; Score 9; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 DIVENPKXL 184
Db 177 DIVENPKXL 185

RESULT 11
A:Accession: A02163
A:Protein: Ig mu heavy chain disease protein (Bot) - human
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999
R:Barikol-Watanabe, S.; Mihaseo, E.; Mihaseo, C.; Barikol, H.U.; Hilschmann, N. Hoppe-Seyler's Z. Physiol. Chem. 365, 105-118, 1984
A:Title: The primary structure of mu-chain-disease protein BOT. Peculiar amino-acid sequence
A:Reference number: A02163; MWID:84184186; PMID:6425189
A:Accession: A02163
A:Molecule type: Protein
A:Residues: 1-391 <BAR>
C:Comment: This protein has no V region homology or CH1 region.
C:Genetics:
A:Gene: GDB:IGHM
A:Cross-references: GDB:120086; OMIM:147020
A:Map position: 1432.33-1432.33
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin; transmembrane prot
F:1-41/Domain: pre-C <VAR>
F:43-391/Domain: Ig mu chain C region, secreted form <IGH>
F:65-137/Domain: immunoglobulin homology <IGH1>
F:175-243/Domain: immunoglobulin homology <IGH2>
F:282-353/Domain: immunoglobulin homology <IGH3>
F:447,210,217,378/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match
Best Local Similarity 3.3%; Score 9; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 TDEVEAEAK 265
Db 99 TDEVEAEAK 107

RESULT 12
A:Accession: E81284
A:Protein: Hypothetical protein Cj1397 [Imported] - *Campylobacter jejuni* (strain NCTC 11168)
C:Species: *Campylobacter jejuni*
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: E81284
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrett, Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp
A:Reference number: A81250; MWID:20150912; PMID:10688204
A:Accession: E81284
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-74 <PAR>
A:Cross-references: GB:A1139078; GB:A111168; NID:96968723; PIDN:CAB73821.1; PID:96968826
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1395; Cj1397

Query Match
Best Local Similarity 2.9%; Score 8; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 TLNELKDG 137
Db 2 TLNELKDG 9

RESULT 13

H75053
molybdopterin converting factor, chain 1 (moad) PAB3357 - Pyrococcus abyssi (strain Orsa)

C/Species: Pyrococcus abyssi
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 17-Feb-2003
C/Accession: H75053

R/Anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A/Reference number: A75001

A/Accession: H75053

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-89 <KAM>

A/Cross-references: GB:AJ248287; GB:AL096836; NID:G5458657; PIDN:CAB50325.1; PID:e151622

A/Experimental source: strain Orsay

C/Genetics:

A/Gene: PAB3357

C/Superfamily: molybdopterin biosynthesis sulfur carrier protein

F/89/Modified site: 1-thioglycine (Gly) #status predicted

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGATI 140

DB 23 ELKDGATI 30

RESULT 14

AE3108
outer membrane lipoprotein Atu4489 (imported) - Agrobacterium tumefaciens (strain C58, D

C/Species: Agrobacterium tumefaciens

C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C/Accession: AE3108

R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A/Reference number: AB2577; MUID:21608550; PMID:11743193

A/Accession: AE3108

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-259 <KUR>

A/Cross-references: GB:AE008689; PIDN:AA45283.1; PID:g17742970; GSPDB:GN00187

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Gene: Atu4489

A/Map position: linear chromosome

C/Superfamily: lipoprotein-28

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGATI 140

RESULT 15

G98178
probable periplasmic protein CJ0772c (imported) - Agrobacterium tumefaciens (strain C58,

C/Species: Agrobacterium tumefaciens

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002

C/Accession: G98178

R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Currello, B.; Goldman,

A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tume

A/Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: G98178

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-259 <KUR>

A/Cross-references: GB:AE007870; PIDN:AAK88953.1; PID:g15158732; GSPDB:GN00170

C/Genetics:

A/Gene: AGR_L_761

A/Map position: linear chromosome

C/Superfamily: lipoprotein-28

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGATI 140

DB 116 ELKDGATI 123

RESULT 16

E64744
probable lipoprotein YaeC - Escherichia coli (strain K-12)

C/Species: Escherichia coli

C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C/Accession: E64744; A47040

R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; MUID:97426617; PMID:9278503

A/Accession: E64744

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-271 <BLAT>

A/Cross-references: GB:AE000129; GB:U00096; NID:g1786395; PIDN:AACT3308.1; PID:g1786396;

A/Experimental source: strain K-12, substrain MG1655

R/Gervais, F.G.; Drapeau, G.R.

J. Bacteriol. 174, 8016-8022, 1992

A/Title: Identification, cloning, and characterization of rcsF, a new regulator gene for

A/Reference number: A47040; MUID:93094132; PMID:1459591

A/Contents: K-12

A/Accession: A47040

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 127-199, 'L', 201-216, 'VSLLKT', 223, 'SPRT' <GER>

A/Cross-references: GB:U04474; NID:g147530; PIDN:AA42507.1; PID:g147531

A/Note: sequence extracted from NCBI backbone (NCBIN:119949; NCBIPI:119950)

C/Genetics:

A/Gene: YaeC

C/Superfamily: lipoprotein-28

Query Match

Best Local Similarity 2.9%; Score 8; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 VGNFTVFP 120

RESULT 17

A83213
Probable ATP-dependent RNA helicase PA3466 (imported) - Pseudomonas aeruginosa (strain P

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C/Accession: A83213

R/Silver, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A/Reference number: A82950; PMID:20437337; PMID:10984043
A/Accession: A83213
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-446 <STO>
A/Cross-references: GB:AE004767; GB:AE004091; NID:G9949604; PINN:AA06854.1; GSPDB:GN001
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA3466

Query Match 2.5%; Score 8; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 149 LARALIL 156
|||||
DB 76 LARALIL 83

RESULT 18
C81400
hypothetical protein Cj0539 [imported] - *Campylobacter jejuni* (strain NCTC 11168)
C/Species: *Campylobacter jejuni*
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C/Accession: C81400
R/Parikh, V.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Church, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanyilet, A.; Whitehead, S.; Barrer
Nature 403, 665-668, 2000
A>Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp
A/Reference number: A81250; PMID:20150912; PMID:10688204
A/Accession: C81400
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-87 <PAR>
A/Cross-references: GB:AL139075; GB:AL111168; NID:G6967817; PINN:CA875175.1; PID:G696800
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: Cj0539

Query Match 2.5%; Score 7; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 LLEKQGL 161
|||||
DB 45 LLEKQGL 51

RESULT 19
B88029
protein F46F5.13 [imported] - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: B88029
R/anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog
A/Reference number: A75000; PMID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_ele
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: B88029
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-104 <STO>
A/Cross-references: GB:chr_II; PINN:AC78191.1; PID:G3886040; GSPDB:GN00020; CESP:F46F5.1
C/Genetics:
A/Gene: F46F5.13
A/Map position: 2

Query Match 2.5%; Score 7; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 108 NNLVIVG 114
|||||
DB 17 NNLVIVG 23

RESULT 20
C90820
hypothetical protein Ecs1531 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: C90820
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gatawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A/Reference number: A99629; PMID:21156231; PMID:11258796
A/Accession: C90820
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-114 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA034954.1; PID:G13360995; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain R1MD 0509952
C/Genetics:
A/Gene: Ecs1531

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 EBAKKQ 267
|||||
DB 54 EBAKKQ 60

RESULT 21
G90851
hypothetical protein Ecs1783 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: G90851
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gatawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A/Reference number: A99629; PMID:21156231; PMID:11258796
A/Accession: G90851
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-114 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA035206.1; PID:G13361248; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain R1MD 0509952
C/Genetics:
A/Gene: Ecs1783

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 EBAKKQ 267
|||||
DB 54 EBAKKQ 60

RESULT 22
C90874
hypothetical protein Ecs1963 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: C90874
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gatawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: C90974
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-114 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA935386.1; PID:g13361428; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs1963

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EAEAKKQ 267
|||||
54 EAEAKKQ 60

Db 54 EAEAKKQ 60

RESULT 23
C90902
hypothetical protein ECs2187 [imported] - *Escherichia coli* (strain O157:H7, substrain R
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: C90902
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: C90902
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-114 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA935610.1; PID:g13361653; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs2187

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EAEAKKQ 267
|||||
54 EAEAKKQ 60

Db 54 EAEAKKQ 60

RESULT 24
D90911
hypothetical protein ECs2260 [imported] - *Escherichia coli* (strain O157:H7, substrain R
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: D90911
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: D90911
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-114 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA935683.1; PID:g13361726; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs2260

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EAEAKKQ 267
|||||
54 EAEAKKQ 60

Db 54 EAEAKKQ 60

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EAEAKKQ 267
|||||
54 EAEAKKQ 60

Db 54 EAEAKKQ 60

RESULT 25
F90971
hypothetical protein ECs2742 [imported] - *Escherichia coli* (strain O157:H7, substrain R
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: F90971
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: F90971
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-114 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA936165.1; PID:g13362210; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs2742

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EAEAKKQ 267
|||||
54 EAEAKKQ 60

Db 54 EAEAKKQ 60

RESULT 26
H91065
hypothetical protein ECs3496 [imported] - *Escherichia coli* (strain O157:H7, substrain R
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: H91065
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: H91065
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-114 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA936919.1; PID:g13362967; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs3496

Query Match 2.5%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EAEAKKQ 267
|||||
54 EAEAKKQ 60

Db 54 EAEAKKQ 60

RESULT 27
D85679
unknown protein encoded by prophage CP-933N [imported] - *Escherichia coli* (strain O157:H
C/Species: *Escherichia coli*
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: D85679
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grothbeck, E.U.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoustis, K.; Apodaca,
Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: D85679
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-114 <STO>
 A:Cross-references: GB:AE005174; NID:G12514706; PIDN:AA655896.1; GSPDB:GN00145; UWGP:Z17
 A:Experimental source: strain O157:H7, substrain EDL933
 A:Genetics:
 A:Gene: Z1795

Query Match 2.5%; Score 7; DB 2; Length 114;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EBAKKQ 267
 |||||
 54 EBAKKQ 60

RESULT 28

unknown protein encoded within prophage CP-933U [imported] - *Escherichia coli* (strain O157:H7)
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: F85815
 R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimatanta, E.; Potamoustis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: F85819
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <STO>
 A:Cross-references: GB:AE005174; NID:G12516122; PIDN:AA657018.1; GSPDB:GN00145; UWGP:Z31
 A:Experimental source: strain O157:H7, substrain EDL933
 A:Genetics:
 A:Gene: Z3105

Query Match 2.5%; Score 7; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EBAKKQ 267
 |||||
 55 EBAKKQ 61

RESULT 29

hypothetical protein Z1351 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85627; H85709
 R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimatanta, E.; Potamoustis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85627
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <STO>
 A:Cross-references: GB:AE005174; NID:G12514192; PIDN:AA655464.1; GSPDB:GN00145; UWGP:Z13
 A:Experimental source: strain O157:H7, substrain EDL933
 A:Accession: H85709
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <STO>
 A:Cross-references: GB:AE005174; NID:G12515016; PIDN:AA656140.1; GSPDB:GN00145; UWGP:Z20
 A:Experimental source: strain O157:H7, substrain EDL933
 A:Genetics:

A:Gene: Z1351; Z2070
 Query Match 2.5%; Score 7; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EBAKKQ 267
 |||||
 55 EBAKKQ 61

RESULT 30

unknown protein encoded within prophage CP-933R [imported] - *Escherichia coli* (strain O157:H7)
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: F85744
 R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimatanta, E.; Potamoustis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: F85744
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-123 <STO>
 A:Cross-references: GB:AE005174; NID:G12515367; PIDN:AA656418.1; GSPDB:GN00145; UWGP:Z23
 A:Experimental source: strain O157:H7, substrain EDL933
 A:Genetics:
 A:Gene: Z2372

Query Match 2.5%; Score 7; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

261 EBAKKQ 267
 |||||
 54 EBAKKQ 60

RESULT 31

hypothetical protein HP0712 - *Helicobacter pylori* (strain 26695)
 C:Species: *Helicobacter pylori*
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
 C:Accession: H64608
 R:Tomb, J.F.; White, O.; Kierlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khajak, H.G.; Glodek, A.; McKenney
 sen, J.D.; Kelley, J.M.; Cotton, M.D.; Weiman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 389, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N
 A>Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: H64608
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-124 <TOM>
 A:Cross-references: GB:AE000584; GB:AE000511; NID:G2313834; PIDN:AA07770.1; PID:G231384

Query Match 2.5%; Score 7; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

71 EFNDYAM 77
 |||||
 8 EFNDYAM 14

RESULT 32

integration host factor, beta subunit XF2437 [imported] - *Xylella fastidiosa* (strain 9a5)
 C:Species: *Xylella fastidiosa*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000

C/Accession: A82559
 R/Anonymous: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:2035717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: A82559
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1126 <SIM>
 A/Cross-references: GB:AE004052; GB:AE003849; NID:g9107617; PIND:AAFP65236.1; GSPDB:GN001
 A/Experimental source: strain 945c
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carrero, D.M.; Carrex, H
 as-Neco, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, W.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
 A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: XF2437
 C/Superfamily: bacterial DNA-binding protein

Query Match 2.5%; Score 7; DB 2; Length 126;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 DDDVLAV 205
 DB 42 DDDVLAV 48

RESULT 33
 A84312
 hypotetical protein Vng1589c [imported] - Halobacterium sp. NRC-1
 C/Species: Halobacterium sp. NRC-1
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C/Accession: A84312
 R/Jing, W.Y.; Kennedy, S.P.; Mahataas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jbld
 Jung, K.H.; Alam, M.; Freiltes, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
 A/Title: Genome sequence of Halobacterium species NRC-1.
 A/Reference number: A84160; MUID:20504483; PMID:11016950
 A/Accession: A84312
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-137 <STO>
 A/Cross-references: GB:AE004437; NID:g10581069; PIND:AAJ19861.1; GSPDB:GN00138
 C/Genetics:
 A/Gene: VNG1589c

Query Match 2.5%; Score 7; DB 2; Length 137;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IALAGS 22
 DB 16 IALAGS 22

RESULT 34
 JE0359
 plasmidocyte spreading peptide precursor protein - Pseudoplusia includens
 C/Species: Pseudoplusia includens

C/Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 07-May-1999
 C/Accession: JE0359
 R/Clark, K.D.; Witherell, A.; Strand, M.R.
 Biochem. Biophys. Res. Commun. 250, 479-485, 1998
 A/Title: Plasmidocyte spreading peptide is encoded by an mRNA differentially expressed in
 A/Reference number: JE0359; MUID:98440825; PMID:9753657
 A/Accession: JE0359
 A/Molecule type: mRNA
 A/Residues: 1-141 <CLA>
 A/Genetics:
 C/Superfamily: paralytic peptide I

Query Match 2.5%; Score 7; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 VFVEDXD 229
 DB 65 VFVEDXD 71

RESULT 35
 S74334
 biotin carboxyl carrier protein - Synechocystis sp. (strain PCC 6803)
 N/Alternate names: hypothetical protein slr0435
 C/Species: Synechocystis sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C/Accession: S74334
 R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
 DNA Res. 3, 109-116, 1996
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 s.
 A/Reference number: S74332; MUID:97061201; PMID:8995231
 A/Accession: S74334
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-154 <KAN>
 A/Cross-references: EMBL:D64001; GB:AB001339; NID:g1001102; PIND:BA10252.1; PID:g100111;
 C/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C/Genetics:
 A/Gene: accB
 A/Start codon: GTG
 C/Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
 F/73-153/Domain: lipoyl/biotin-binding homology <LPB>

Query Match 2.5%; Score 7; DB 2; Length 154;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 AEVAGQV 58
 DB 126 AEVAGQV 132

RESULT 36
 I67751
 dopamine receptor D5 - green monkey (fragment)
 C/Species: Cercopithecus aethiops (green monkey, grivet)
 C/Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 13-Aug-1999
 C/Accession: I67751
 R/Nguyen, T.; Bard, J.A.; Jin, H.; Tarnusio, D.; Ward, D.C.; Kennedy, J.L.; Weinschenk, R.
 Gene 109, 211-218, 1991
 A/Title: Human dopamine D5 receptor pseudogenes.
 A/Reference number: I53655; MUID:92112045; PMID:1765268
 A/Accession: I67751
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-156 <R5>
 A/Cross-references: GB:M77187; NID:g342376; PIND:AAA36923.1; PID:g342377
 C/Genetics:

A:Gene: DRD5
C:Superfamily: vertebrate rhodopsin

Query Match 2.5%; Score 7; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 VAAVAG 56
|||||
DB 25 VAAVAG 31

RESULT 37

C97438

hypothetical protein AGR_C.1160 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C:Accession: C97438

R:Gooder, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman,

A: Liu, F.; Woliam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97459; PMID:21608551; PMID:11743194

A:Accession: C97438

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-188 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86460.1; PID:G15155604; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C.1160

A:Map position: circular chromosome

Query Match

Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VAAARAD 199
|||||
DB 71 VAAARAD 77

RESULT 38

AF2656

conserved hypothetical protein Atu0652 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AF2656

R:Wood, D.W.; Seubal, U.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AF2656

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-188 <KUR>

A:Cross-references: GB:AE008688; PIDN:AA14166.1; PID:G17733011; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu0652

A:Map position: circular chromosome

Query Match

Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 VAAARAD 199
|||||
DB 71 VAAARAD 77

RESULT 39

JQ2185

coat protein - apple chlorotic leaf spot virus (isolate P-205)

N:Alternate names: ORF 3 protein

C:Species: apple chlorotic leaf spot virus

C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000

C:Accession: JQ2185

R:Sato, K.; Yoshikawa, N.; Takahashi, T.

J. Gen. Virol. 74, 1927-1931, 1993

A:Title: Complete nucleotide sequence of the genome of an apple isolate of apple chlorot

A:Reference number: JQ2183; PMID:93389448; PMID:8376968

A:Accession: JQ2185

A:Molecule type: mRNA

A:Residues: 1-193 <SNAP>

A:Cross-references: GB:D14996; NID:9434059; PIDN:BA03643.1; PID:9453242

C:Superfamily: apple chlorotic leaf spot virus coat protein

C:Keywords: coat protein

QY 168 TNLPSTT 174
|||||
DB 126 TNLPSTT 132

RESULT 40

AC3629

3-oxoacyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.-) [imported] - Brucella melitensis

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AC3629

R:DeVeecho, V.G.; Kapatal, V.; Redhar, R.J.; Patra, G.; Muier, C.; Los, T.; Ivanova, J

.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessc

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AC3629

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-196 <KUR>

A:Cross-references: GB:AE008918; PIDN:AA154198.1; PID:G17985166; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMET10956

A:Map position: II

C:Superfamily: deaF protein

C:Keywords: carbon-carbon lyase; carboxy-lyase

QY 193 VAAARAD 199
|||||
DB 167 VAAARAD 173

Search completed: June 16, 2004, 11:17:36
Job time : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 11:14:18 ; Search time 23 Seconds
(without alignments)
619,512 Million cell updates/sec

Title: US-10-018-672-2
Perfect score: 276
Sequence: 1 MNFGKINGICATASGIALG.....TDEVAEAKKQFDVIGKGM 276

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PCUTS_COMB.pep.*
6: /cgn2_6/prodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	20.7	118	US-09-540-236-2991	Sequence 2991, Ap
2	16	5.8	279	US-09-543-681A-6534	Sequence 6534, Ap
3	10	3.6	286	US-09-328-352-6702	Sequence 6702, Ap
4	9	3.3	285	US-09-489-039A-10641	Sequence 10641, A
5	9	3.3	291	US-09-107-532A-7234	Sequence 7234, Ap
6	9	3.3	296	US-09-328-352-6924	Sequence 6924, Ap
7	8	2.9	269	US-09-107-532A-6443	Sequence 6443, Ap
8	8	2.9	273	US-09-252-991A-19629	Sequence 19629, A
9	8	2.9	279	US-09-543-681A-4548	Sequence 4548, Ap
10	8	2.9	279	US-09-543-681A-6914	Sequence 6914, Ap
11	8	2.9	289	US-09-489-039A-12337	Sequence 12337, A
12	8	2.9	465	US-08-252-991A-17635	Sequence 17635, A
13	7	2.5	78	US-08-858-207A-422	Sequence 422, App
14	7	2.5	116	US-09-134-001C-3331	Sequence 3331, Ap
15	7	2.5	159	US-09-963-791-18	Sequence 18, Appl
16	7	2.5	161	US-08-572-447C-11	Sequence 11, Appl
17	7	2.5	161	US-09-267-747-11	Sequence 11, Appl
18	7	2.5	199	PCT-US94-02889-2	Sequence 2, Appl
19	7	2.5	226	US-08-572-447C-15	Sequence 15, Appl
20	7	2.5	226	US-09-267-747-15	Sequence 15, Appl
21	7	2.5	274	US-09-543-681A-4708	Sequence 4708, Ap
22	7	2.5	275	US-09-489-039A-7722	Sequence 7722, Ap
23	7	2.5	298	US-09-434-354-49	Sequence 49, Appl
24	7	2.5	298	US-09-434-354-48	Sequence 49, Appl
25	7	2.5	302	US-09-134-001C-3171	Sequence 3171, Ap
26	7	2.5	306	US-09-252-991A-18673	Sequence 18673, A
27	7	2.5	309	US-09-522-714-10	Sequence 10, Appl

28	7	2.5	310	US-09-963-791-8	Sequence 8, Appl
29	7	2.5	317	US-09-963-791-16	Sequence 16, Appl
30	7	2.5	326	US-09-489-039A-9512	Sequence 9512, Ap
31	7	2.5	342	US-09-540-236-2111	Sequence 2111, Ap
32	7	2.5	347	US-09-543-681A-5961	Sequence 5961, Ap
33	7	2.5	351	US-09-252-991A-30094	Sequence 30094, A
34	7	2.5	356	US-09-963-791-20	Sequence 20, Appl
35	7	2.5	361	US-09-543-681A-5390	Sequence 5390, Ap
36	7	2.5	419	US-08-305-127B-4	Sequence 4, Appl
37	7	2.5	438	US-09-963-791-22	Sequence 22, Appl
38	7	2.5	447	US-09-252-991A-20116	Sequence 20116, A
39	7	2.5	468	US-09-963-791-6	Sequence 6, Appl
40	7	2.5	475	US-07-686-591-4	Sequence 4, Appl
41	7	2.5	475	US-07-970-715-4	Sequence 4, Appl
42	7	2.5	477	US-07-791-936A-2	Sequence 2, Appl
43	7	2.5	477	US-08-383-781B-2	Sequence 2, Appl
44	7	2.5	477	US-07-969-267B-2	Sequence 2, Appl
45	7	2.5	477	US-07-969-267B-3	Sequence 3, Appl
46	7	2.5	477	US-09-168-510-2	Sequence 2, Appl
47	7	2.5	477	US-09-168-510-3	Sequence 3, Appl
48	7	2.5	506	US-09-134-001C-4383	Sequence 4383, Ap
49	7	2.5	507	US-09-963-791-10	Sequence 10, Appl
50	7	2.5	581	US-09-543-681A-4472	Sequence 4472, Ap
51	7	2.5	589	US-09-963-791-12	Sequence 12, Appl
52	7	2.5	652	US-08-987-119-2	Sequence 2, Appl
53	7	2.5	690	US-09-336-115C-22	Sequence 22, Appl
54	7	2.5	691	US-09-336-115C-12	Sequence 12, Appl
55	7	2.5	757	US-09-963-791-24	Sequence 24, Appl
56	7	2.5	859	US-09-369-364A-5	Sequence 5, Appl
57	7	2.5	886	US-09-252-991A-26999	Sequence 26999, A
58	7	2.5	908	US-09-963-791-2	Sequence 2, Appl
59	7	2.5	982	US-09-252-991A-30580	Sequence 30580, A
60	7	2.5	1482	US-09-410-551B-21	Sequence 21, Appl
61	7	2.5	1488	US-09-410-551B-17	Sequence 17, Appl
62	7	2.5	1509	US-09-410-551B-23	Sequence 23, Appl
63	7	2.5	1517	US-09-410-551B-19	Sequence 19, Appl
64	7	2.5	1969	US-09-418-710-72	Sequence 72, Appl
65	7	2.5	1972	US-09-418-710-21	Sequence 21, Appl
66	7	2.5	6396	US-09-410-551B-72	Sequence 72, Appl
67	6	2.2	14	US-08-165-038-5	Sequence 5, Appl
68	6	2.2	14	US-08-876-781-5	Sequence 5, Appl
69	6	2.2	14	US-09-567-003C-2	Sequence 2, Appl
70	6	2.2	17	US-08-290-448A-55	Sequence 55, Appl
71	6	2.2	17	US-08-290-448A-55	Sequence 55, Appl
72	6	2.2	17	US-08-175-069A-55	Sequence 55, Appl
73	6	2.2	17	US-08-461-939B-55	Sequence 55, Appl
74	6	2.2	17	US-08-464-000-55	Sequence 55, Appl
75	6	2.2	20	US-09-009-993-193	Sequence 193, App
76	6	2.2	25	US-09-515-356-3	Sequence 3, Appl
77	6	2.2	25	US-09-515-356-4	Sequence 4, Appl
78	6	2.2	27	US-09-053-717A-69	Sequence 69, Appl
79	6	2.2	27	US-09-085-761A-74	Sequence 74, Appl
80	6	2.2	27	US-09-017-689A-1	Sequence 1, Appl
81	6	2.2	30	US-08-290-448A-17	Sequence 17, Appl
82	6	2.2	30	US-08-290-448A-17	Sequence 17, Appl
83	6	2.2	30	US-08-175-069A-17	Sequence 17, Appl
84	6	2.2	30	US-08-461-939B-17	Sequence 17, Appl
85	6	2.2	30	US-08-464-000-17	Sequence 17, Appl
86	6	2.2	34	US-08-129-089-5	Sequence 5, Appl
87	6	2.2	34	US-08-129-089-8	Sequence 8, Appl
88	6	2.2	34	US-08-129-089-9	Sequence 9, Appl
89	6	2.2	34	US-08-129-089-10	Sequence 10, Appl
90	6	2.2	34	US-08-129-089-11	Sequence 11, Appl
91	6	2.2	34	US-08-129-089-12	Sequence 12, Appl
92	6	2.2	34	US-08-454-444-1	Sequence 1, Appl
93	6	2.2	34	US-08-454-444-4	Sequence 4, Appl
94	6	2.2	34	US-08-454-444-5	Sequence 5, Appl
95	6	2.2	34	US-08-454-444-6	Sequence 6, Appl
96	6	2.2	34	US-08-454-444-7	Sequence 7, Appl
97	6	2.2	34	US-08-454-444-8	Sequence 8, Appl
98	6	2.2	34	PCT-US93-01135-5	Sequence 5, Appl
99	6	2.2	34	PCT-US93-01135-8	Sequence 8, Appl
100	6	2.2	34	PCT-US93-01135-9	Sequence 9, Appl

101	6	2.2	34	5	PCT-US93-01135-10	Sequence 10, Appl	174	6	2.2	166	4	US-09-339-913B-85	Sequence 85, Appl
102	6	2.2	34	5	PCT-US93-01135-11	Sequence 11, Appl	175	6	2.2	166	4	US-09-339-904A-75	Sequence 75, Appl
103	6	2.2	34	5	PCT-US93-01135-12	Sequence 12, Appl	176	6	2.2	166	4	US-09-339-904A-76	Sequence 76, Appl
104	6	2.2	42	1	US-08-230-448A-13	Sequence 13, Appl	177	6	2.2	166	4	US-09-339-904A-77	Sequence 77, Appl
105	6	2.2	42	1	US-08-230-448A-13	Sequence 13, Appl	178	6	2.2	166	4	US-09-339-904A-79	Sequence 79, Appl
106	6	2.2	42	1	US-08-175-069A-13	Sequence 13, Appl	179	6	2.2	166	4	US-09-339-904A-81	Sequence 81, Appl
107	6	2.2	42	1	US-08-461-939B-13	Sequence 13, Appl	180	6	2.2	166	4	US-09-339-904A-82	Sequence 82, Appl
108	6	2.2	42	4	US-08-464-000-13	Sequence 13, Appl	181	6	2.2	166	4	US-09-339-904A-84	Sequence 84, Appl
109	6	2.2	44	4	US-08-858-207A-280	Sequence 280, Appl	182	6	2.2	166	4	US-09-339-904A-85	Sequence 85, Appl
110	6	2.2	44	1	US-08-230-448A-12	Sequence 12, Appl	183	6	2.2	166	4	US-09-339-904A-86	Sequence 86, Appl
111	6	2.2	45	1	US-08-290-448A-12	Sequence 12, Appl	184	6	2.2	166	4	US-09-324-542-109	Sequence 109, Appl
112	6	2.2	45	1	US-08-175-069A-12	Sequence 12, Appl	185	6	2.2	166	4	US-08-769-062B-75	Sequence 75, Appl
113	6	2.2	45	4	US-08-461-939B-12	Sequence 12, Appl	186	6	2.2	166	4	US-08-769-062B-76	Sequence 76, Appl
114	6	2.2	45	4	US-08-464-000-12	Sequence 12, Appl	187	6	2.2	166	4	US-08-769-062B-77	Sequence 77, Appl
115	6	2.2	54	1	US-08-165-038-1	Sequence 1, Appl	188	6	2.2	166	4	US-08-769-062B-81	Sequence 81, Appl
116	6	2.2	54	1	US-08-621-081A-1	Sequence 1, Appl	189	6	2.2	166	4	US-08-769-062B-82	Sequence 82, Appl
117	6	2.2	54	2	US-08-876-781-1	Sequence 1, Appl	190	6	2.2	166	4	US-08-769-062B-84	Sequence 84, Appl
118	6	2.2	54	4	US-08-936-165A-320	Sequence 320, Appl	191	6	2.2	166	4	US-08-769-062B-85	Sequence 85, Appl
119	6	2.2	58	4	US-09-621-976-6622	Sequence 6622, Ap	192	6	2.2	166	4	US-08-769-062B-86	Sequence 86, Appl
120	6	2.2	61	4	US-08-858-207A-343	Sequence 343, Ap	193	6	2.2	166	4	US-09-344-002B-75	Sequence 75, Appl
121	6	2.2	63	4	US-09-252-991A-19534	Sequence 19534, A	194	6	2.2	166	4	US-09-344-002B-76	Sequence 76, Appl
122	6	2.2	66	4	US-09-227-357-654	Sequence 654, Ap	195	6	2.2	166	4	US-09-344-002B-77	Sequence 77, Appl
123	6	2.2	70	4	US-09-489-039A-12899	Sequence 12899, A	196	6	2.2	166	4	US-09-344-002B-79	Sequence 79, Appl
124	6	2.2	82	4	US-09-252-991A-18536	Sequence 18536, A	197	6	2.2	166	4	US-09-344-002B-81	Sequence 81, Appl
125	6	2.2	82	4	US-09-489-039A-14014	Sequence 14014, A	198	6	2.2	166	4	US-09-344-002B-82	Sequence 82, Appl
126	6	2.2	84	4	US-09-252-991A-18487	Sequence 18487, A	199	6	2.2	166	4	US-09-344-002B-84	Sequence 84, Appl
127	6	2.2	92	4	US-09-621-976-5000	Sequence 5000, Ap	200	6	2.2	166	4	US-09-344-002B-85	Sequence 85, Appl
128	6	2.2	99	4	US-09-540-236-2992	Sequence 2992, Ap	201	6	2.2	166	4	US-09-344-002B-86	Sequence 86, Appl
129	6	2.2	100	4	US-09-732-210-109	Sequence 709, Ap	202	6	2.2	166	4	US-09-205-466-109	Sequence 109, Appl
130	6	2.2	107	4	US-09-328-352-5562	Sequence 5562, Ap	203	6	2.2	166	4	US-09-559-555C-75	Sequence 75, Appl
131	6	2.2	111	1	US-08-466-886-38	Sequence 38, Appl	204	6	2.2	166	4	US-09-559-555C-76	Sequence 76, Appl
132	6	2.2	111	3	US-08-469-617-38	Sequence 38, Appl	205	6	2.2	166	4	US-09-559-555C-77	Sequence 77, Appl
133	6	2.2	114	4	US-09-489-039A-8548	Sequence 8548, Ap	206	6	2.2	166	4	US-09-559-555C-79	Sequence 79, Appl
134	6	2.2	116	4	US-09-732-210-1691	Sequence 1691, Ap	207	6	2.2	166	4	US-09-559-555C-81	Sequence 81, Appl
135	6	2.2	117	1	US-07-942-245-24	Sequence 24, Appl	208	6	2.2	166	4	US-09-559-555C-82	Sequence 82, Appl
136	6	2.2	117	4	US-09-732-210-732	Sequence 732, Appl	209	6	2.2	166	4	US-09-559-555C-84	Sequence 84, Appl
137	6	2.2	118	4	US-09-252-991A-147	Sequence 147, Ap	210	6	2.2	166	4	US-09-559-555C-85	Sequence 85, Appl
138	6	2.2	119	4	US-09-732-210-147	Sequence 29926, A	211	6	2.2	166	4	US-09-559-555C-86	Sequence 86, Appl
139	6	2.2	122	4	US-09-732-210-685	Sequence 685, Ap	212	6	2.2	166	4	US-09-559-555C-87	Sequence 87, Appl
140	6	2.2	124	3	US-08-513-974B-349	Sequence 349, Ap	213	6	2.2	166	4	US-09-252-991A-32785	Sequence 32785, A
141	6	2.2	127	4	US-09-252-991A-21888	Sequence 21888, A	214	6	2.2	166	4	US-09-693-350-75	Sequence 75, Appl
142	6	2.2	128	4	US-09-621-976-5844	Sequence 5844, Ap	215	6	2.2	166	4	US-09-693-350-76	Sequence 76, Appl
143	6	2.2	131	4	US-09-247-155-175	Sequence 175, Ap	216	6	2.2	166	4	US-09-693-350-77	Sequence 77, Appl
144	6	2.2	140	4	US-09-134-000C-6522	Sequence 6522, Ap	217	6	2.2	166	4	US-09-693-350-79	Sequence 79, Appl
145	6	2.2	141	4	US-09-107-532A-4712	Sequence 4712, Ap	218	6	2.2	166	4	US-09-693-350-81	Sequence 81, Appl
146	6	2.2	146	4	US-09-543-681A-7317	Sequence 7317, Ap	219	6	2.2	166	4	US-09-693-350-82	Sequence 82, Appl
147	6	2.2	147	4	US-09-134-001C-3024	Sequence 3024, Ap	220	6	2.2	166	4	US-09-693-350-84	Sequence 84, Appl
148	6	2.2	151	4	US-09-732-210-1674	Sequence 1674, Ap	221	6	2.2	166	4	US-09-693-350-85	Sequence 85, Appl
149	6	2.2	151	4	US-09-489-039A-9773	Sequence 9773, Ap	222	6	2.2	166	4	US-09-693-350-86	Sequence 86, Appl
150	6	2.2	152	4	US-09-252-991A-26218	Sequence 26218, A	223	6	2.2	166	4	US-09-693-389-75	Sequence 75, Appl
151	6	2.2	153	4	US-09-732-210-1672	Sequence 1672, Ap	224	6	2.2	166	4	US-09-693-389-76	Sequence 76, Appl
152	6	2.2	155	4	US-09-370-838-114	Sequence 114, Ap	225	6	2.2	166	4	US-09-693-389-77	Sequence 77, Appl
153	6	2.2	156	1	PCT-US94-06447-5	Sequence 5, Appl	226	6	2.2	166	4	US-09-693-389-79	Sequence 79, Appl
154	6	2.2	156	5	PCT-US94-06447-5	Sequence 5, Appl	227	6	2.2	166	4	US-09-693-389-81	Sequence 81, Appl
155	6	2.2	160	4	US-09-252-991A-25771	Sequence 25771, A	228	6	2.2	166	4	US-09-693-389-82	Sequence 82, Appl
156	6	2.2	161	4	US-09-252-991A-23802	Sequence 23802, A	229	6	2.2	166	4	US-09-693-389-84	Sequence 84, Appl
157	6	2.2	166	1	US-08-288-671-1	Sequence 1, Appl	230	6	2.2	166	4	US-09-693-389-85	Sequence 85, Appl
158	6	2.2	166	1	US-08-489-066A-16	Sequence 16, Appl	231	6	2.2	166	4	US-06-236-204C-66	Sequence 66, Appl
159	6	2.2	166	2	US-08-997-080-109	Sequence 109, Appl	232	6	2.2	166	4	US-06-236-204C-72	Sequence 72, Appl
160	6	2.2	166	2	US-08-875-802-4	Sequence 4, Appl	233	6	2.2	166	4	US-06-256-204C-73	Sequence 73, Appl
161	6	2.2	166	2	US-08-997-362-109	Sequence 109, Ap	234	6	2.2	166	4	US-06-256-204C-75	Sequence 75, Appl
162	6	2.2	166	3	US-08-489-072A-16	Sequence 16, Appl	235	6	2.2	166	4	US-06-256-204C-77	Sequence 77, Appl
163	6	2.2	166	3	US-09-085-855-109	Sequence 109, Ap	236	6	2.2	166	4	US-06-256-204C-78	Sequence 78, Appl
164	6	2.2	166	3	US-09-305-723C-4	Sequence 4, Appl	237	6	2.2	166	4	US-06-256-204C-80	Sequence 80, Appl
165	6	2.2	166	3	US-08-489-071A-16	Sequence 16, Appl	238	6	2.2	166	4	US-06-256-204C-83	Sequence 83, Appl
166	6	2.2	166	4	US-09-339-913B-75	Sequence 75, Appl	239	6	2.2	166	4	US-06-256-204C-85	Sequence 85, Appl
167	6	2.2	166	4	US-09-339-913B-76	Sequence 76, Appl	240	6	2.2	166	4	US-09-559-671A-75	Sequence 75, Appl
168	6	2.2	166	4	US-09-339-913B-77	Sequence 77, Appl	241	6	2.2	166	4	US-09-559-671A-76	Sequence 76, Appl
169	6	2.2	166	4	US-09-339-913B-79	Sequence 79, Appl	242	6	2.2	166	4	US-09-559-671A-77	Sequence 79, Appl
170	6	2.2	166	4	US-09-339-913B-81	Sequence 81, Appl	243	6	2.2	166	4	US-09-559-671A-79	Sequence 79, Appl
171	6	2.2	166	4	US-09-339-913B-82	Sequence 82, Appl	244	6	2.2	166	4	US-09-559-671A-81	Sequence 81, Appl
172	6	2.2	166	4	US-09-339-913B-84	Sequence 84, Appl	245	6	2.2	166	4	US-09-559-671A-82	Sequence 82, Appl
173	6	2.2	166	4	US-09-339-913B-85	Sequence 85, Appl	246	6	2.2	166	4	US-09-559-671A-84	Sequence 84, Appl

247	6	2.2	166	4	US-09-559-671A-85	Sequence 85, Appl	320	6	2.2	189	3	US-09-206-935-13	Sequence 13, Appl
248	6	2.2	166	4	US-09-559-671A-86	Sequence 86, Appl	321	6	2.2	189	3	US-09-206-935-14	Sequence 14, Appl
249	6	2.2	166	4	US-09-339-926A-75	Sequence 75, Appl	322	6	2.2	189	3	US-09-206-935-15	Sequence 15, Appl
250	6	2.2	166	4	US-09-339-926A-76	Sequence 76, Appl	323	6	2.2	189	3	US-09-206-935-17	Sequence 17, Appl
251	6	2.2	166	4	US-09-339-926A-77	Sequence 77, Appl	324	6	2.2	189	3	US-09-206-935-18	Sequence 18, Appl
252	6	2.2	166	4	US-09-339-926A-79	Sequence 79, Appl	325	6	2.2	189	3	US-09-206-935-19	Sequence 19, Appl
253	6	2.2	166	4	US-09-339-926A-81	Sequence 81, Appl	326	6	2.2	189	4	US-08-489-071A-2	Sequence 2, Appl
254	6	2.2	166	4	US-09-339-926A-82	Sequence 82, Appl	327	6	2.2	189	4	US-08-489-071A-3	Sequence 3, Appl
255	6	2.2	166	4	US-09-339-926A-84	Sequence 84, Appl	328	6	2.2	189	4	US-09-206-936-10	Sequence 10, Appl
256	6	2.2	166	4	US-09-339-926A-85	Sequence 85, Appl	329	6	2.2	189	4	US-09-206-936-13	Sequence 13, Appl
257	6	2.2	166	4	US-09-339-926A-86	Sequence 86, Appl	330	6	2.2	189	4	US-09-206-936-15	Sequence 15, Appl
258	6	2.2	166	4	US-09-744-754C-9	Sequence 9, Appl	331	6	2.2	189	4	US-09-206-936-17	Sequence 17, Appl
259	6	2.2	166	4	US-09-744-754C-11	Sequence 11, Appl	332	6	2.2	189	4	US-09-206-936-18	Sequence 18, Appl
260	6	2.2	166	4	US-09-744-754C-30	Sequence 30, Appl	333	6	2.2	189	4	US-09-206-936-19	Sequence 19, Appl
261	6	2.2	166	4	US-09-744-754C-32	Sequence 32, Appl	334	6	2.2	189	4	US-09-487-792-7	Sequence 7, Appl
262	6	2.2	166	4	US-09-744-754C-34	Sequence 34, Appl	335	6	2.2	189	4	US-09-487-792-8	Sequence 8, Appl
263	6	2.2	166	4	US-09-744-754C-36	Sequence 36, Appl	336	6	2.2	189	4	US-09-908-594-7	Sequence 7, Appl
264	6	2.2	166	4	US-09-744-754C-38	Sequence 38, Appl	337	6	2.2	189	4	US-09-908-594-8	Sequence 8, Appl
265	6	2.2	166	4	US-09-744-754C-40	Sequence 40, Appl	338	6	2.2	189	4	US-07-145-002B-4	Sequence 4, Appl
266	6	2.2	166	4	US-09-744-754C-42	Sequence 42, Appl	339	6	2.2	189	4	US-07-145-002B-16	Sequence 16, Appl
267	6	2.2	167	2	US-08-875-802-2	Sequence 2, Appl	340	6	2.2	189	4	US-07-145-002B-12	Sequence 12, Appl
268	6	2.2	167	2	US-08-875-802-3	Sequence 3, Appl	341	6	2.2	189	4	US-07-145-002B-18	Sequence 18, Appl
269	6	2.2	167	3	US-09-305-723C-2	Sequence 2, Appl	342	6	2.2	189	4	US-07-145-002B-19	Sequence 19, Appl
270	6	2.2	167	3	US-09-305-723C-3	Sequence 3, Appl	343	6	2.2	189	4	US-07-145-002B-22	Sequence 22, Appl
271	6	2.2	167	4	US-07-145-002B-46	Sequence 46, Appl	344	6	2.2	189	4	US-07-145-002B-30	Sequence 30, Appl
272	6	2.2	167	4	US-07-145-002B-47	Sequence 47, Appl	345	6	2.2	189	4	US-07-145-002B-32	Sequence 32, Appl
273	6	2.2	167	4	US-07-145-002B-49	Sequence 49, Appl	346	6	2.2	189	4	US-07-145-002B-37	Sequence 37, Appl
274	6	2.2	167	4	US-07-145-002B-51	Sequence 51, Appl	347	6	2.2	189	4	US-06-256-204C-6	Sequence 6, Appl
275	6	2.2	167	4	US-07-145-002B-52	Sequence 52, Appl	348	6	2.2	189	4	US-06-256-204C-12	Sequence 12, Appl
276	6	2.2	167	4	US-07-145-002B-54	Sequence 54, Appl	349	6	2.2	189	4	US-06-256-204C-18	Sequence 18, Appl
277	6	2.2	167	4	US-07-145-002B-57	Sequence 57, Appl	350	6	2.2	189	4	US-06-256-204C-30	Sequence 30, Appl
278	6	2.2	167	4	US-07-145-002B-59	Sequence 59, Appl	351	6	2.2	189	4	US-06-256-204C-32	Sequence 32, Appl
279	6	2.2	167	4	US-06-256-204C-46	Sequence 46, Appl	352	6	2.2	189	4	US-06-256-204C-37	Sequence 37, Appl
280	6	2.2	167	4	US-06-256-204C-47	Sequence 47, Appl	353	6	2.2	189	4	US-06-256-204C-50	Sequence 50, Appl
281	6	2.2	167	4	US-06-256-204C-49	Sequence 49, Appl	354	6	2.2	189	4	US-06-256-204C-52	Sequence 52, Appl
282	6	2.2	167	4	US-06-256-204C-51	Sequence 51, Appl	355	6	2.2	189	4	US-06-256-204C-57	Sequence 57, Appl
283	6	2.2	167	4	US-06-256-204C-52	Sequence 52, Appl	356	6	2.2	191	4	US-09-252-991A-26824	Sequence 26824, Ap
284	6	2.2	167	4	US-06-256-204C-54	Sequence 54, Appl	357	6	2.2	203	4	US-08-956-171E-5207	Sequence 5207, Ap
285	6	2.2	167	4	US-06-256-204C-57	Sequence 57, Appl	358	6	2.2	203	4	US-09-489-039A-1975	Sequence 1975, Ap
286	6	2.2	167	4	US-06-256-204C-59	Sequence 59, Appl	359	6	2.2	204	4	US-09-134-001C-2975	Sequence 2975, Ap
287	6	2.2	169	4	US-07-145-002B-34	Sequence 34, Appl	360	6	2.2	205	4	US-09-252-991A-24225	Sequence 24225, A
288	6	2.2	169	4	US-07-145-002B-38	Sequence 38, Appl	361	6	2.2	209	4	US-09-199-637A-172	Sequence 172, Ap
289	6	2.2	169	4	US-06-256-204C-34	Sequence 34, Appl	362	6	2.2	210	4	US-09-543-681A-8215	Sequence 8215, Ap
290	6	2.2	169	4	US-06-256-204C-38	Sequence 38, Appl	363	6	2.2	212	4	US-09-328-352-5416	Sequence 5416, Ap
291	6	2.2	169	4	US-09-134-000C-3567	Sequence 3567, Ap	364	6	2.2	214	4	US-08-165-038-4	Sequence 4, Appl
292	6	2.2	170	4	US-09-252-991A-27069	Sequence 27069, A	365	6	2.2	214	4	US-08-876-781A-4	Sequence 4, Appl
293	6	2.2	170	4	US-09-107-532A-5197	Sequence 5197, Ap	366	6	2.2	213	1	US-08-876-781A-4	Sequence 4, Appl
294	6	2.2	172	4	US-09-328-352-6306	Sequence 6306, Ap	367	6	2.2	213	2	US-09-252-991A-28039	Sequence 28039, A
295	6	2.2	175	4	US-09-651-941-13	Sequence 13, Appl	368	6	2.2	213	4	US-09-134-000C-4384	Sequence 4384, Ap
296	6	2.2	175	4	US-09-955-597-13	Sequence 13, Appl	369	6	2.2	213	4	US-09-252-991A-19712	Sequence 19712, A
297	6	2.2	179	4	US-09-489-039A-8913	Sequence 8913, Ap	370	6	2.2	214	4	US-08-914-375C-29	Sequence 29, Appl
298	6	2.2	182	4	US-02-107-532A-6641	Sequence 6641, Ap	371	6	2.2	215	4	US-09-291-170A-6	Sequence 6, Appl
299	6	2.2	182	4	US-09-050-759-12	Sequence 12, Appl	372	6	2.2	215	4	US-08-724-884-6	Sequence 6, Appl
300	6	2.2	185	4	US-09-198-452A-106	Sequence 106, Ap	373	6	2.2	215	4	US-08-621-081A-13	Sequence 13, Appl
301	6	2.2	186	4	US-09-489-039A-12559	Sequence 12559, A	374	6	2.2	215	4	US-08-621-081A-21	Sequence 21, Appl
302	6	2.2	188	6	5510472-8	Patent No. 5510472	375	6	2.2	217	1	US-08-621-081A-22	Sequence 22, Appl
303	6	2.2	189	1	US-08-026-758-1	Sequence 1, Appl	376	6	2.2	217	1	US-08-621-081A-29	Sequence 29, Appl
304	6	2.2	189	1	US-08-026-758-9	Sequence 9, Appl	377	6	2.2	217	1	US-08-621-081A-33	Sequence 33, Appl
305	6	2.2	189	1	US-08-026-758-10	Sequence 10, Appl	378	6	2.2	217	1	US-08-621-081A-44	Sequence 44, Appl
306	6	2.2	189	1	US-08-026-758-11	Sequence 11, Appl	379	6	2.2	217	1	US-08-621-081A-45	Sequence 45, Appl
307	6	2.2	189	1	US-08-026-758-12	Sequence 12, Appl	380	6	2.2	217	1	US-08-621-081A-25	Sequence 25, Appl
308	6	2.2	189	1	US-08-026-758-13	Sequence 13, Appl	381	6	2.2	217	1	US-08-621-081A-57	Sequence 57, Appl
309	6	2.2	189	1	US-08-026-758-14	Sequence 14, Appl	382	6	2.2	217	1	US-08-621-081A-58	Sequence 58, Appl
310	6	2.2	189	1	US-08-026-758-15	Sequence 15, Appl	383	6	2.2	217	1	US-08-621-081A-29	Sequence 29, Appl
311	6	2.2	189	1	US-08-026-758-16	Sequence 16, Appl	384	6	2.2	217	1	US-08-262-451B-4	Sequence 24723, A
312	6	2.2	189	1	US-08-026-758-17	Sequence 17, Appl	385	6	2.2	219	1	US-08-621-081A-16	Sequence 16, Appl
313	6	2.2	189	1	US-08-026-758-18	Sequence 18, Appl	386	6	2.2	219	1	US-08-621-081A-24	Sequence 24, Appl
314	6	2.2	189	1	US-08-026-758-20	Sequence 20, Appl	387	6	2.2	219	1	US-08-621-081A-19	Sequence 19, Appl
315	6	2.2	189	1	US-08-489-066A-2	Sequence 2, Appl	388	6	2.2	221	1	US-08-621-081A-24	Sequence 24, Appl
316	6	2.2	189	2	US-08-489-066A-3	Sequence 3, Appl	389	6	2.2	221	4	US-09-252-991A-22003	Sequence 22003, A
317	6	2.2	189	3	US-08-489-072A-2	Sequence 2, Appl	390	6	2.2	223	4	US-09-252-991A-32061	Sequence 32061, A
318	6	2.2	189	3	US-08-489-072A-3	Sequence 3, Appl	391	6	2.2	224	4	US-09-393-634-9	Sequence 9, Appl
319	6	2.2	189	3	US-09-206-935-10	Sequence 10, Appl	392	6	2.2	224	4		

393	6	2.2	225	4	US-09-540-236-1659	Sequence 3659, Ap	466	6	2.2	310	2	US-08-337-358-42	Sequence 42, Appl
394	6	2.2	226	4	US-09-489-039A-7662	Sequence 7662, Ap	467	6	2.2	310	5	PCT-US95-07537A-42	Sequence 42, Appl
395	6	2.2	226	4	US-09-540-236-1787	Sequence 3787, Ap	468	6	2.2	310	5	PCT-US95-07537A-42	Sequence 42, Appl
396	6	2.2	228	3	US-09-248-335-70	Sequence 70, Appl	469	6	2.2	311	4	US-09-252-991A-17395	Sequence 17395, A
397	6	2.2	229	4	US-09-328-352-7445	Sequence 7445, Ap	470	6	2.2	311	4	US-09-107-532A-7665	Sequence 7665, Ap
398	6	2.2	229	4	US-09-543-681A-6872	Sequence 6872, Ap	471	6	2.2	311	4	US-09-489-039A-9441	Sequence 9441, Ap
399	6	2.2	232	1	US-08-638-911A-33	Sequence 33, Appl	472	6	2.2	311	4	US-09-489-039A-13109	Sequence 13109, A
400	6	2.2	236	1	US-08-442-063A-42	Sequence 42, Appl	473	6	2.2	311	4	US-09-489-039A-13269	Sequence 13269, A
401	6	2.2	237	4	US-09-489-039A-12899	Sequence 12899, A	474	6	2.2	313	4	US-09-489-039A-8623	Sequence 8623, Ap
402	6	2.2	239	4	US-09-328-932A-65	Sequence 65, Appl	475	6	2.2	316	4	US-09-252-991A-29147	Sequence 29147, A
403	6	2.2	244	4	US-09-107-532A-5393	Sequence 5393, Ap	476	6	2.2	316	4	US-09-328-352-5588	Sequence 5588, Ap
404	6	2.2	249	2	US-08-797-689-18	Sequence 18, Appl	477	6	2.2	316	4	US-09-134-000C-5598	Sequence 5598, Ap
405	6	2.2	249	2	US-09-984-186-18	Sequence 18, Appl	478	6	2.2	322	4	US-09-252-991A-22816	Sequence 22816, A
406	6	2.2	251	4	US-09-253-991A-24493	Sequence 24493, A	479	6	2.2	324	4	US-09-489-039A-7632	Sequence 7632, A
407	6	2.2	252	4	US-09-253-991A-29152	Sequence 29152, A	480	6	2.2	329	4	US-09-252-991A-27332	Sequence 27332, A
408	6	2.2	253	4	US-09-489-039A-10070	Sequence 10070, A	481	6	2.2	330	4	US-09-489-039A-13373	Sequence 13373, Ap
409	6	2.2	255	4	US-09-489-039A-11759	Sequence 11759, A	482	6	2.2	330	4	US-09-543-681A-5560	Sequence 5560, Ap
410	6	2.2	256	4	US-09-328-352-7574	Sequence 7574, Ap	483	6	2.2	331	2	US-08-997-060-182	Sequence 182, Ap
411	6	2.2	257	4	US-09-071-035-32	Sequence 32, Appl	484	6	2.2	331	2	US-08-997-362-182	Sequence 182, Ap
412	6	2.2	258	4	US-09-489-039A-12038	Sequence 12038, A	485	6	2.2	331	4	US-09-095-855-182	Sequence 182, Ap
413	6	2.2	260	3	US-09-961-083-32	Sequence 32, Appl	486	6	2.2	331	4	US-09-324-542-182	Sequence 182, Ap
414	6	2.2	260	3	US-09-536-784-32	Sequence 32, Appl	487	6	2.2	331	4	US-09-205-446-182	Sequence 446, Ap
415	6	2.2	261	4	US-09-634-238-420	Sequence 420, Ap	488	6	2.2	332	4	US-09-543-681A-27332	Sequence 27332, A
416	6	2.2	263	4	US-09-489-039A-8239	Sequence 8239, Ap	489	6	2.2	333	1	US-09-442-063A-27	Sequence 27, Appl
417	6	2.2	264	4	US-09-328-352-6699	Sequence 6699, Ap	490	6	2.2	333	1	US-09-328-352-7378	Sequence 7378, Ap
418	6	2.2	266	4	US-09-543-681A-4676	Sequence 4676, Ap	491	6	2.2	334	3	US-08-975-762-28	Sequence 28, Appl
419	6	2.2	270	4	US-09-252-991A-24772	Sequence 24772, A	492	6	2.2	334	3	US-08-821-324-28	Sequence 28, Appl
420	6	2.2	271	4	US-09-252-991A-17679	Sequence 17679, A	493	6	2.2	334	3	US-09-295-058-28	Sequence 28, Appl
421	6	2.2	271	4	US-09-252-991A-20200	Sequence 20200, A	494	6	2.2	334	3	US-09-106-582-28	Sequence 28, Appl
422	6	2.2	272	4	US-09-489-039A-10410	Sequence 10410, A	495	6	2.2	334	4	US-09-159-469-28	Sequence 28, Appl
423	6	2.2	275	4	US-09-647-324A-26	Sequence 26, Appl	496	6	2.2	334	4	US-09-693-542-28	Sequence 28, Appl
424	6	2.2	277	4	US-09-134-000C-6416	Sequence 6416, Ap	497	6	2.2	338	3	US-09-922-957-1	Sequence 1, Appl
425	6	2.2	277	4	US-09-071-035-30	Sequence 30, Appl	498	6	2.2	338	3	US-08-922-957-3	Sequence 3, Appl
426	6	2.2	278	4	US-09-134-000C-4965	Sequence 4965, Ap	499	6	2.2	339	3	US-09-134-001C-3913	Sequence 3913, Ap
427	6	2.2	279	4	US-09-134-001C-4878	Sequence 4878, Ap	500	6	2.2	340	1	US-07-828-700-8	Sequence 8, Appl
428	6	2.2	279	4	US-09-489-039A-13210	Sequence 13210, A	501	6	2.2	340	1	US-09-489-039A-10379	Sequence 10379, A
429	6	2.2	280	4	US-09-956-171E-5185	Sequence 5185, Ap	502	6	2.2	341	4	US-09-543-681A-4755	Sequence 4755, Ap
430	6	2.2	281	4	US-09-543-681A-4336	Sequence 4336, Ap	503	6	2.2	342	1	US-08-272-919-2	Sequence 2, Appl
431	6	2.2	282	1	US-08-442-063A-45	Sequence 45, Appl	504	6	2.2	342	1	US-08-615-916-2	Sequence 2, Appl
432	6	2.2	282	4	US-09-252-991A-25948	Sequence 25948, A	505	6	2.2	342	4	US-09-134-001C-4190	Sequence 4190, Ap
433	6	2.2	283	4	US-09-107-532A-6360	Sequence 6360, Ap	506	6	2.2	342	4	US-09-253-991A-25420	Sequence 25420, A
434	6	2.2	285	4	US-09-489-039A-12529	Sequence 12529, A	507	6	2.2	342	5	PCT-US95-08542-2	Sequence 2, Appl
435	6	2.2	286	4	US-09-252-991A-20535	Sequence 20535, A	508	6	2.2	346	4	US-09-328-352-6625	Sequence 6625, Ap
436	6	2.2	286	4	US-09-543-681A-7127	Sequence 7127, Ap	509	6	2.2	347	4	US-09-252-991A-27330	Sequence 27330, A
437	6	2.2	288	4	US-09-134-001C-4005	Sequence 4005, Ap	510	6	2.2	347	4	US-09-328-352-7050	Sequence 7050, Ap
438	6	2.2	289	4	US-09-134-000C-4722	Sequence 4722, Ap	511	6	2.2	347	4	US-09-215-418-5	Sequence 5, Appl
439	6	2.2	289	4	US-09-540-236-2018	Sequence 2018, Ap	512	6	2.2	348	4	US-09-489-039A-8632	Sequence 8632, Ap
440	6	2.2	291	4	US-09-252-991A-28572	Sequence 28572, A	513	6	2.2	350	1	US-07-759-568-3	Sequence 3, Appl
441	6	2.2	292	4	US-09-134-001C-3479	Sequence 3479, Ap	514	6	2.2	350	3	US-08-430-286A-8	Sequence 8, Appl
442	6	2.2	292	4	US-09-134-000C-3587	Sequence 3587, Ap	515	6	2.2	350	3	US-09-252-991A-30112	Sequence 30112, A
443	6	2.2	293	4	US-09-252-991A-23350	Sequence 23350, A	516	6	2.2	351	4	US-09-489-039A-7599	Sequence 7599, Ap
444	6	2.2	294	4	US-09-134-000C-4724	Sequence 4724, Ap	517	6	2.2	352	4	US-09-489-039A-8236	Sequence 8236, Ap
445	6	2.2	294	4	US-09-134-000C-4825	Sequence 4825, Ap	518	6	2.2	353	6	5340334-4	Patent No. 5340334
446	6	2.2	296	4	US-09-543-681A-4665	Sequence 4665, Ap	519	6	2.2	356	4	US-09-252-991A-29888	Sequence 29888, A
447	6	2.2	297	4	US-09-252-991A-28075	Sequence 28075, A	520	6	2.2	357	3	US-08-809-286B-3	Sequence 3, Appl
448	6	2.2	298	3	US-09-922-957-4	Sequence 4, Appl	521	6	2.2	357	4	US-09-543-681A-4670	Sequence 4670, A
449	6	2.2	298	4	US-09-107-532A-6664	Sequence 6664, Ap	522	6	2.2	358	4	US-09-107-532A-4143	Sequence 4143, Ap
450	6	2.2	300	4	US-09-328-352-6778	Sequence 6778, Ap	523	6	2.2	359	1	US-08-303-338-4	Sequence 4, Appl
451	6	2.2	302	4	US-09-252-991A-24353	Sequence 24353, A	524	6	2.2	359	3	US-08-458-834-4	Sequence 4, Appl
452	6	2.2	303	3	US-08-158-735A-14	Sequence 14, Appl	525	6	2.2	364	4	US-09-134-000C-3748	Sequence 3748, Ap
453	6	2.2	303	3	US-08-158-735A-15	Sequence 15, Appl	526	6	2.2	365	4	US-09-252-991A-20904	Sequence 20904, A
454	6	2.2	303	4	US-09-105-258-818	Sequence 818, Appl	527	6	2.2	366	4	US-08-545-573A-1	Sequence 1, Appl
455	6	2.2	305	4	US-09-328-352-7028	Sequence 7028, Ap	528	6	2.2	370	3	US-08-467-023-97	Sequence 97, Appl
456	6	2.2	305	4	US-09-252-991A-22096	Sequence 22096, A	529	6	2.2	370	4	US-09-252-991A-23112	Sequence 23112, A
457	6	2.2	306	4	US-09-160-975A-2	Sequence 2, Appl	530	6	2.2	370	4	US-09-134-000C-4746	Sequence 4746, Ap
458	6	2.2	306	4	US-09-107-532A-5464	Sequence 5464, Ap	531	6	2.2	372	1	US-08-597-236-8	Sequence 8, Appl
459	6	2.2	307	1	US-08-442-063A-48	Sequence 48, Appl	532	6	2.2	372	1	US-08-746-682A-8	Sequence 8, Appl
460	6	2.2	307	4	US-09-634-238-245	Sequence 245, Appl	533	6	2.2	372	4	US-09-328-352-5436	Sequence 5436, Ap
461	6	2.2	307	4	US-09-107-532A-4626	Sequence 4626, Ap	534	6	2.2	374	4	US-09-252-991A-24475	Sequence 24475, A
462	6	2.2	308	4	US-09-107-532A-5553	Sequence 5553, Ap	535	6	2.2	375	1	US-08-121-714-5	Sequence 5, Appl
463	6	2.2	308	4	US-09-489-039A-7802	Sequence 7802, Ap	536	6	2.2	375	1	US-08-477-108A-5	Sequence 5, Appl
464	6	2.2	309	4	US-09-134-001C-4952	Sequence 4952, Ap	537	6	2.2	375	2	US-08-477-112-5	Sequence 5, Appl
465	6	2.2	310	1	US-08-433-783-42	Sequence 42, Appl	538	6	2.2	375	4	US-09-328-352-6078	Sequence 6078, Ap

539	6	2.2	375	4	US-09-489-039A-11560	Sequence 11560, A	612	6	2.2	437	4	US-09-134-000C-3874	Sequence 3874, Ap
540	6	2.2	375	5	PCT-US93-08322-5	Sequence 5, Appl1	613	6	2.2	438	4	US-09-540-236-2955	Sequence 2955, Ap
541	6	2.2	376	4	US-09-198-452A-995	Sequence 995, Appl	614	6	2.2	439	1	US-08-333-386-10	Sequence 10, Appl
542	6	2.2	376	4	US-09-252-991A-19776	Sequence 19776, A	615	6	2.2	439	1	US-08-463-694-10	Sequence 10, Appl
543	6	2.2	383	1	US-08-290-448A-78	Sequence 78, Appl	616	6	2.2	439	1	US-08-694-501-10	Sequence 10, Appl
544	6	2.2	383	1	US-08-290-448A-78	Sequence 78, Appl	617	6	2.2	440	2	US-08-484-200-4	Sequence 4, Appl
545	6	2.2	383	1	US-08-175-069A-78	Sequence 78, Appl	618	6	2.2	440	4	US-09-198-452A-44	Sequence 44, Appl
546	6	2.2	383	4	US-08-461-939B-78	Sequence 78, Appl	619	6	2.2	441	4	US-09-107-532A-4625	Sequence 4625, Ap
547	6	2.2	383	4	US-08-461-939B-78	Sequence 78, Appl	620	6	2.2	441	4	US-09-107-532A-4625	Sequence 4625, Ap
548	6	2.2	384	2	US-08-833-226-2	Sequence 2, Appl1	621	6	2.2	445	4	US-09-134-001C-1948	Sequence 6, Appl1
549	6	2.2	386	4	US-08-545-573A-2	Sequence 2, Appl1	622	6	2.2	447	3	US-08-508-761B-6	Sequence 6, Appl1
550	6	2.2	386	4	US-08-545-573A-39	Sequence 39, Appl	623	6	2.2	454	4	US-09-252-991A-30693	Sequence 30693, A
551	6	2.2	387	1	US-08-290-448A-72	Sequence 72, Appl	624	6	2.2	456	1	US-08-624-125-20	Sequence 20, Appl
552	6	2.2	387	1	US-08-290-448A-72	Sequence 72, Appl	625	6	2.2	456	4	US-08-937-155-20	Sequence 20, Appl
553	6	2.2	387	1	US-08-175-069A-72	Sequence 72, Appl	626	6	2.2	456	4	US-09-253-991A-11884	Sequence 20, Appl
554	6	2.2	387	4	US-08-461-939B-72	Sequence 72, Appl	627	6	2.2	458	4	US-09-253-991A-22614	Sequence 22614, A
555	6	2.2	387	4	US-08-461-939B-72	Sequence 72, Appl	628	6	2.2	463	3	US-08-792-295-1	Sequence 1, Appl1
556	6	2.2	387	1	US-08-290-448A-59	Sequence 59, Appl	629	6	2.2	463	3	US-09-076-432-1	Sequence 1, Appl1
557	6	2.2	391	1	US-08-290-448A-59	Sequence 59, Appl	630	6	2.2	463	4	US-09-198-452A-115	Sequence 115, App
558	6	2.2	391	1	US-08-175-069A-59	Sequence 59, Appl	631	6	2.2	467	4	US-09-134-001C-3579	Sequence 3579, Ap
559	6	2.2	391	4	US-08-461-939B-59	Sequence 59, Appl	632	6	2.2	469	4	US-09-489-039A-13565	Sequence 13565, A
560	6	2.2	391	4	US-08-464-000-59	Sequence 59, Appl	633	6	2.2	471	3	US-09-160-494-6	Sequence 6, Appl1
561	6	2.2	393	4	US-09-107-532A-4627	Sequence 4627, Ap	634	6	2.2	475	4	US-09-328-352-4942	Sequence 4942, Ap
562	6	2.2	396	4	US-08-236-3802	Sequence 3802, Ap	635	6	2.2	479	4	US-09-253-991A-23608	Sequence 23608, A
563	6	2.2	397	1	US-08-290-448A-76	Sequence 76, Appl	636	6	2.2	482	4	US-09-252-991A-23559	Sequence 23559, A
564	6	2.2	397	1	US-08-290-448A-76	Sequence 76, Appl	637	6	2.2	482	4	US-09-252-991A-28368	Sequence 28368, A
565	6	2.2	397	1	US-08-175-069A-76	Sequence 76, Appl	638	6	2.2	483	4	US-09-336-115C-20	Sequence 20, Appl
566	6	2.2	397	4	US-08-461-939B-76	Sequence 76, Appl	639	6	2.2	484	1	US-08-597-236-2	Sequence 2, Appl1
567	6	2.2	397	4	US-08-464-000-76	Sequence 76, Appl	640	6	2.2	484	1	US-08-746-682A-2	Sequence 2, Appl1
568	6	2.2	398	1	US-08-290-448A-74	Sequence 74, Appl	641	6	2.2	485	4	US-09-253-991A-21841	Sequence 21841, A
569	6	2.2	398	1	US-08-290-448A-74	Sequence 74, Appl	642	6	2.2	486	4	US-09-489-039A-8181	Sequence 8181, Ap
570	6	2.2	398	1	US-08-175-069A-74	Sequence 74, Appl	643	6	2.2	487	4	US-09-252-991A-31302	Sequence 31302, A
571	6	2.2	398	4	US-08-461-939B-74	Sequence 74, Appl	644	6	2.2	489	4	US-09-107-532A-4711	Sequence 4711, Ap
572	6	2.2	398	4	US-08-464-000-74	Sequence 74, Appl	645	6	2.2	494	3	US-08-984-618-3	Sequence 3, Appl1
573	6	2.2	399	4	US-09-252-991A-31028	Sequence 31028, A	646	6	2.2	498	4	US-09-252-991A-28182	Sequence 28182, A
574	6	2.2	400	4	US-09-255-258-822	Sequence 822, App	647	6	2.2	498	4	US-09-323-998B-57	Sequence 57, Appl
575	6	2.2	401	1	US-08-368-803-7	Sequence 7, Appl1	648	6	2.2	500	4	US-09-499-302A-5	Sequence 5, Appl
576	6	2.2	402	4	US-09-489-039A-9618	Sequence 9618, Ap	649	6	2.2	500	4	US-09-323-998B-58	Sequence 58, Appl
577	6	2.2	404	4	US-09-198-452A-718	Sequence 718, App	650	6	2.2	500	4	US-09-323-998B-59	Sequence 59, Appl
578	6	2.2	404	4	US-09-489-039A-10542	Sequence 10542, A	651	6	2.2	502	3	US-09-382-256-12	Sequence 12, Appl
579	6	2.2	405	2	US-08-222-719-2	Sequence 2, Appl1	652	6	2.2	502	3	US-09-395-115-12	Sequence 12, Appl
580	6	2.2	405	2	US-08-470-925-2	Sequence 2, Appl1	653	6	2.2	502	4	US-08-436-265-12	Sequence 12, Appl
581	6	2.2	405	2	US-08-471-613-2	Sequence 2, Appl1	654	6	2.2	502	4	US-09-679-197-12	Sequence 12, Appl
582	6	2.2	405	4	US-09-107-532A-6414	Sequence 6414, Ap	655	6	2.2	502	4	US-09-499-302A-2	Sequence 2, Appl1
583	6	2.2	405	5	PCT-US93-10443-2	Sequence 2, Appl1	656	6	2.2	502	4	US-09-499-302A-4	Sequence 4, Appl1
584	6	2.2	407	4	US-09-252-991A-27349	Sequence 27349, A	657	6	2.2	502	4	US-09-134-000C-6324	Sequence 6324, Ap
585	6	2.2	407	4	US-09-252-991A-29531	Sequence 29531, A	658	6	2.2	502	4	US-09-323-998B-56	Sequence 56, Appl
586	6	2.2	409	4	US-09-613-303-55	Sequence 55, Appl	659	6	2.2	503	2	US-08-481-337A-2	Sequence 2, Appl1
587	6	2.2	409	4	US-10-267-311-55	Sequence 55, Appl	660	6	2.2	503	2	US-08-696-268B-2	Sequence 2, Appl1
588	6	2.2	410	4	US-09-252-991A-30590	Sequence 30590, A	661	6	2.2	503	3	US-09-382-256-2	Sequence 2, Appl1
589	6	2.2	410	4	US-09-198-452A-50	Sequence 50, Appl	662	6	2.2	503	3	US-09-395-115-2	Sequence 2, Appl1
590	6	2.2	412	4	US-09-252-991A-17613	Sequence 17613, A	663	6	2.2	503	4	US-08-679-265-2	Sequence 2, Appl1
591	6	2.2	412	4	US-09-252-991A-125085	Sequence 125085, A	664	6	2.2	503	4	US-09-679-187-2	Sequence 2, Appl1
592	6	2.2	412	4	US-09-134-000C-6666	Sequence 6666, Ap	665	6	2.2	503	4	US-08-448-371A-2	Sequence 2, Appl1
593	6	2.2	413	4	US-09-489-039A-13838	Sequence 13838, A	666	6	2.2	503	5	PCT-US94-11328A-4	Sequence 4, Appl1
594	6	2.2	414	4	US-09-489-039A-12570	Sequence 12570, A	667	6	2.2	503	5	PCT-US95-05467-2	Sequence 2, Appl1
595	6	2.2	417	4	US-09-489-039A-12011	Sequence 12011, A	668	6	2.2	505	1	US-08-149-105-16	Sequence 16, Appl
596	6	2.2	422	2	US-08-663-566A-5	Sequence 5, Appl1	669	6	2.2	505	1	US-08-317-847-14	Sequence 14, Appl
597	6	2.2	422	2	US-08-023-610-5	Sequence 5, Appl1	670	6	2.2	505	4	US-09-543-681A-7151	Sequence 7151, Ap
598	6	2.2	422	2	US-08-288-065A-5	Sequence 5, Appl1	671	6	2.2	505	4	US-09-489-039A-8637	Sequence 8637, Ap
599	6	2.2	422	2	US-08-362-240A-5	Sequence 5, Appl1	672	6	2.2	507	4	US-09-091-097-14	Sequence 10, Appl
600	6	2.2	422	2	US-08-804-372A-3	Sequence 3, Appl1	673	6	2.2	507	4	US-09-091-097-14	Sequence 10, Appl
601	6	2.2	422	5	PCT-US95-10245-5	Sequence 5, Appl1	674	6	2.2	507	4	US-09-252-991A-16830	Sequence 16830, A
602	6	2.2	423	4	US-09-134-001C-4717	Sequence 4717, Ap	675	6	2.2	509	1	US-08-149-105-14	Sequence 14, Appl
603	6	2.2	423	4	US-09-252-991A-21128	Sequence 21128, A	676	6	2.2	509	1	US-08-317-847-14	Sequence 14, Appl
604	6	2.2	424	4	US-09-252-991A-18970	Sequence 18970, A	677	6	2.2	509	2	US-08-481-337A-4	Sequence 4, Appl1
605	6	2.2	425	4	US-09-252-991A-30735	Sequence 30735, A	678	6	2.2	509	3	US-09-382-256-4	Sequence 4, Appl1
606	6	2.2	426	4	US-09-489-039A-9881	Sequence 9881, Ap	679	6	2.2	509	3	US-09-395-115-4	Sequence 4, Appl1
607	6	2.2	429	1	US-08-745-977-4	Sequence 4, Appl1	680	6	2.2	509	3	US-08-123-934A-6	Sequence 6, Appl1
608	6	2.2	429	3	US-09-040-699A-4	Sequence 4, Appl1	681	6	2.2	509	4	US-08-436-265-4	Sequence 4, Appl1
609	6	2.2	431	4	US-09-543-681A-7623	Sequence 7623, Ap	682	6	2.2	509	4	US-09-679-187-4	Sequence 4, Appl1
610	6	2.2	433	4	US-09-252-991A-28824	Sequence 28824, A	683	6	2.2	509	4	US-09-874-626-6	Sequence 6, Appl1
611	6	2.2	435	4	US-09-252-991A-28507	Sequence 28507, A	684	6	2.2	509	4	US-08-448-371A-4	Sequence 4, Appl1

685	6	2.2	509	5	PCT-US94-10080-6	Sequence 6, Appl1	758	6	2.2	594	4	US-09-570-383-2	Sequence 2, Appl1
686	6	2.2	509	5	PCT-US95-05467-4	Sequence 4, Appl1	759	6	2.2	594	4	US-09-395-336-2	Sequence 2, Appl1
687	6	2.2	510	4	US-09-252-991A-19584	Sequence 15684, A	760	6	2.2	602	4	US-09-540-236-263	Sequence 3,633, Ap
688	6	2.2	511	4	US-09-134-000C-5362	Sequence 5362, Ap	761	6	2.2	603	4	US-09-543-681A-5587	Sequence 5587, Ap
689	6	2.2	513	4	US-09-489-039A-14224	Sequence 14224, A	762	6	2.2	604	4	US-09-328-352-4437	Sequence 6437, Ap
690	6	2.2	519	4	US-09-252-991A-30815	Sequence 30815, A	763	6	2.2	608	4	US-09-134-000C-4096	Sequence 4096, Ap
691	6	2.2	522	4	US-09-543-681A-7135	Sequence 7135, Ap	764	6	2.2	609	4	US-09-252-991A-32547	Sequence 32547, A
692	6	2.2	525	4	US-09-676-166A-3	Sequence 3, Appl1	765	6	2.2	610	4	US-09-252-991A-27740	Sequence 27740, A
693	6	2.2	525	4	US-09-252-991A-20491	Sequence 20491, A	766	6	2.2	613	4	US-09-252-991A-21411	Sequence 21411, A
694	6	2.2	526	4	US-09-548-938A-10	Sequence 10, Appl1	767	6	2.2	614	3	US-09-017-706-10	Sequence 9, Appl1
695	6	2.2	526	4	US-09-489-039A-8223	Sequence 8223, Ap	768	6	2.2	614	3	US-09-017-706-10	Sequence 11, Appl1
696	6	2.2	527	2	US-08-659-939-4	Sequence 4, Appl1	769	6	2.2	614	3	US-09-017-706-12	Sequence 12, Appl1
697	6	2.2	527	2	US-08-850-041-4	Sequence 4, Appl1	770	6	2.2	614	3	US-09-017-706-13	Sequence 13, Appl1
698	6	2.2	527	3	US-08-896-500-4	Sequence 4, Appl1	771	6	2.2	614	3	US-09-017-706-14	Sequence 14, Appl1
699	6	2.2	527	3	US-08-476-397-4	Sequence 4, Appl1	772	6	2.2	614	3	US-09-567-003C-22	Sequence 22, Appl1
700	6	2.2	527	3	US-08-973-720-4	Sequence 4, Appl1	773	6	2.2	622	1	US-08-547-137-1	Sequence 1, Appl1
701	6	2.2	527	3	US-08-923-558-2	Sequence 2, Appl1	774	6	2.2	622	1	US-08-957-940-1	Sequence 1, Appl1
702	6	2.2	527	3	US-09-262-927-4	Sequence 4, Appl1	775	6	2.2	631	4	US-09-252-991A-18000	Sequence 18000, A
703	6	2.2	527	4	US-09-486-553-2	Sequence 2, Appl1	776	6	2.2	631	4	US-09-640-958-10	Sequence 10, Appl1
704	6	2.2	527	4	US-09-570-383-4	Sequence 4, Appl1	777	6	2.2	634	4	US-09-252-991A-33100	Sequence 33100, A
705	6	2.2	527	4	US-09-395-336-4	Sequence 4, Appl1	778	6	2.2	635	4	US-09-129-452A-128	Sequence 128, Ap
706	6	2.2	529	4	US-09-149-476-732	Sequence 732, Ap	779	6	2.2	636	4	US-09-134-000C-4594	Sequence 4594, Ap
707	6	2.2	529	4	US-09-328-352-6406	Sequence 6406, Ap	780	6	2.2	637	4	US-09-336-115C-10	Sequence 14, Appl1
708	6	2.2	530	4	US-08-975-762-73	Sequence 73, Appl1	781	6	2.2	640	4	US-09-252-991A-24554	Sequence 24554, A
709	6	2.2	530	3	US-09-295-028-73	Sequence 73, Appl1	782	6	2.2	652	4	US-09-252-991A-17333	Sequence 17333, A
710	6	2.2	530	4	US-09-106-582-73	Sequence 73, Appl1	783	6	2.2	659	4	US-09-134-001C-4303	Sequence 4303, Ap
711	6	2.2	530	4	US-09-159-469-73	Sequence 73, Appl1	784	6	2.2	661	4	US-09-556-916-26	Sequence 26, Appl1
712	6	2.2	530	4	US-09-693-542-73	Sequence 73, Appl1	785	6	2.2	672	4	US-09-556-916-26	Sequence 28, Appl1
713	6	2.2	541	4	US-09-252-991A-20524	Sequence 20524, A	786	6	2.2	672	4	US-09-336-115C-10	Sequence 10, Appl1
714	6	2.2	541	4	US-09-252-991A-32947	Sequence 32947, A	787	6	2.2	674	4	US-09-129-452A-128	Sequence 18107, A
715	6	2.2	542	4	US-09-215-694-9	Sequence 9, Appl1	788	6	2.2	674	4	US-09-134-000C-4004	Sequence 4004, Ap
716	6	2.2	544	4	US-09-252-991A-19874	Sequence 19874, A	789	6	2.2	681	4	US-09-252-991A-29071	Sequence 29071, A
717	6	2.2	546	4	US-09-066-047-2	Sequence 2, Appl1	790	6	2.2	685	6	5194595-19	Patent No. 5194595
718	6	2.2	546	4	US-09-489-039A-9627	Sequence 9627, Ap	791	6	2.2	684	4	US-09-252-991A-24802	Sequence 24802, A
719	6	2.2	548	2	US-08-457-822-31	Sequence 31, Appl1	792	6	2.2	684	4	US-09-328-352-5781	Sequence 5781, Ap
720	6	2.2	548	3	US-08-432-697-31	Sequence 31, Appl1	793	6	2.2	684	4	US-09-328-352-5781	Sequence 29072, A
721	6	2.2	548	3	US-08-466-248-31	Sequence 31, Appl1	794	6	2.2	700	4	US-09-252-991A-29072	Sequence 8176, Ap
722	6	2.2	548	4	US-09-149-476-469	Sequence 469, Ap	795	6	2.2	702	4	US-09-328-352-5781	Sequence 3591, Ap
723	6	2.2	554	4	US-09-564-805-236	Sequence 236, Ap	796	6	2.2	703	4	US-09-540-226-3691	Sequence 4, Appl1
724	6	2.2	560	1	US-08-052-404-23	Sequence 23, Appl1	797	6	2.2	707	4	US-09-021-560-4	Sequence 2, Appl1
725	6	2.2	560	1	US-08-479-156-23	Sequence 23, Appl1	798	6	2.2	708	4	US-09-336-115C-2	Sequence 5094, Ap
726	6	2.2	561	1	US-08-052-404-24	Sequence 24, Appl1	799	6	2.2	715	4	US-09-134-000C-5094	Sequence 2, Appl1
727	6	2.2	561	1	US-08-479-156-24	Sequence 24, Appl1	800	6	2.2	721	4	US-09-853-533A-2	Sequence 2, Appl1
728	6	2.2	562	3	US-09-227-717-2	Sequence 2, Appl1	801	6	2.2	726	4	US-09-184-748-2	Sequence 2, Appl1
729	6	2.2	562	4	US-09-173-151A-22	Sequence 22, Appl1	802	6	2.2	733	4	US-09-328-352-5781	Sequence 2411, A
730	6	2.2	567	4	US-09-198-452A-994	Sequence 994, Ap	803	6	2.2	741	4	US-09-252-991A-20098	Sequence 20098, A
731	6	2.2	573	4	US-09-252-991A-24488	Sequence 24488, A	804	6	2.2	744	4	US-09-021-560-2	Sequence 6, Appl1
732	6	2.2	573	4	US-09-134-000C-5258	Sequence 5258, Ap	805	6	2.2	745	4	US-09-336-115C-6	Sequence 19, Appl1
733	6	2.2	574	4	US-09-252-991A-18912	Sequence 18912, A	806	6	2.2	748	3	US-09-061-764A-19	Sequence 26975, A
734	6	2.2	574	4	US-09-252-991A-26497	Sequence 26497, A	807	6	2.2	750	4	US-09-252-991A-26975	Sequence 2, Appl1
735	6	2.2	574	4	5223254-2	Patent No. 5223254	808	6	2.2	751	4	US-09-468-415-2	Sequence 10222, A
736	6	2.2	577	4	US-09-489-039A-9575	Sequence 9575, A	809	6	2.2	754	4	US-08-547-197-2	Sequence 2, Appl1
737	6	2.2	580	4	US-09-252-991A-22943	Sequence 22943, A	810	6	2.2	760	1	US-08-957-940-2	Sequence 31816, A
738	6	2.2	585	4	US-09-134-001C-4914	Sequence 4914, Ap	811	6	2.2	760	1	US-09-252-991A-31816	Sequence 13, Appl1
739	6	2.2	590	3	US-08-975-762-64	Sequence 64, Appl1	812	6	2.2	764	4	US-09-252-991A-31816	Sequence 11, Appl1
740	6	2.2	590	3	US-08-295-028-64	Sequence 64, Appl1	813	6	2.2	769	4	US-09-788-657-15	Sequence 13, Appl1
741	6	2.2	590	4	US-09-106-582-64	Sequence 64, Appl1	814	6	2.2	770	2	US-08-209-521-13	Sequence 123, Ap
742	6	2.2	590	4	US-09-159-469-64	Sequence 64, Appl1	815	6	2.2	770	3	US-08-961-810-123	Sequence 123, Ap
743	6	2.2	590	4	US-09-693-542-64	Sequence 64, Appl1	816	6	2.2	770	3	US-08-352-902D-123	Sequence 123, Ap
744	6	2.2	593	4	US-09-328-352-4866	Sequence 4866, Ap	817	6	2.2	770	3	US-09-252-991A-20455	Sequence 20455, A
745	6	2.2	594	2	US-08-659-939-2	Sequence 2, Appl1	818	6	2.2	771	4	US-09-252-991A-31403	Sequence 31403, A
746	6	2.2	594	2	US-08-850-041-2	Sequence 2, Appl1	819	6	2.2	773	4	US-09-252-991A-29487	Sequence 29487, A
747	6	2.2	594	2	US-08-457-822-31	Sequence 31, Appl1	820	6	2.2	774	4	US-08-186-228-50	Sequence 50, Appl1
748	6	2.2	594	2	US-08-838-189D-6	Sequence 6, Appl1	821	6	2.2	780	1	US-08-333-638-14	Sequence 40, Appl1
749	6	2.2	594	3	US-08-852-344D-6	Sequence 6, Appl1	822	6	2.2	780	1	US-08-333-638-14	Sequence 50, Appl1
750	6	2.2	594	3	US-08-896-500-2	Sequence 2, Appl1	823	6	2.2	781	2	US-08-676-631-3	Sequence 3, Appl1
751	6	2.2	594	3	US-08-476-397-2	Sequence 2, Appl1	824	6	2.2	781	2	US-08-359-316A-3	Sequence 3, Appl1
752	6	2.2	594	3	US-08-973-720-2	Sequence 2, Appl1	825	6	2.2	801	3	US-09-147-226-4	Sequence 2, Appl1
753	6	2.2	594	3	US-08-344-639E-6	Sequence 6, Appl1	826	6	2.2	802	4	US-09-522-474-4	Sequence 4, Appl1
754	6	2.2	594	3	US-08-262-927-2	Sequence 2, Appl1	827	6	2.2	802	4		
755	6	2.2	594	3	US-08-467-669A-6	Sequence 6, Appl1	828	6	2.2				
756	6	2.2	594	3	US-08-467-661A-6	Sequence 6, Appl1	829	6	2.2				
757	6	2.2	594	3	US-08-001-554A-6	Sequence 6, Appl1	830	6	2.2				

831	6	2.2	804	1	US-08-785-428-2	Sequence 2, Appl	904	6	2.2	1647	4	US-09-535-008-2	Sequence 2, Appl
832	6	2.2	804	2	US-08-996-797-2	Sequence 2, Appl	905	6	2.2	1649	4	US-09-535-008-2	Sequence 2, Appl
833	6	2.2	813	4	US-09-328-352-7421	Sequence 7421, Ap	906	6	2.2	1650	4	US-09-535-008-11	Sequence 71, Appl
834	6	2.2	819	4	US-09-328-352-5342	Sequence 5342, Ap	907	6	2.2	1678	4	US-09-535-008-55	Sequence 69, Appl
835	6	2.2	826	5	US-07-638-431-2	Sequence 2, Appl	908	6	2.2	1679	4	US-09-535-008-55	Sequence 65, Appl
836	6	2.2	826	5	PCT-US92-00018-2	Sequence 2, Appl	909	6	2.2	1681	4	US-09-535-008-77	Sequence 77, Appl
837	6	2.2	826	5	US-09-491-356C-20	Sequence 19, Appl	910	6	2.2	1682	4	US-09-535-008-73	Sequence 73, Appl
838	6	2.2	832	4	US-09-585-858-19	Sequence 3, Appl	911	6	2.2	1690	4	US-09-595-684B-39	Sequence 39, Appl
839	6	2.2	860	1	US-08-117-362-3	Sequence 3, Appl	912	6	2.2	1784	3	US-09-040-778-2	Sequence 2, Appl
840	6	2.2	860	1	US-08-486-924A-3	Sequence 3, Appl	913	6	2.2	1784	3	US-08-652-46A-2	Sequence 2, Appl
841	6	2.2	860	1	US-08-486-924A-3	Sequence 3, Appl	914	6	2.2	2101	1	US-08-466-350-4	Sequence 4, Appl
842	6	2.2	871	4	US-09-352-991A-33057	Sequence 33057, A	915	6	2.2	2101	1	US-08-470-950-4	Sequence 4, Appl
843	6	2.2	873	4	US-09-352-991A-30504	Sequence 30504, A	916	6	2.2	2101	1	US-08-467-781-4	Sequence 4, Appl
844	6	2.2	879	1	US-08-306-546C-2	Sequence 2, Appl	917	6	2.2	2101	1	US-08-195-487-4	Sequence 4, Appl
845	6	2.2	880	2	US-08-530-524A-2	Sequence 2, Appl	918	6	2.2	2101	1	US-08-483-924-4	Sequence 4, Appl
846	6	2.2	880	2	US-09-252-991A-25046	Sequence 25046, A	919	6	2.2	2101	2	US-09-452-284-1	Sequence 4, Appl
847	6	2.2	883	4	US-09-489-039A-12755	Sequence 12755, A	920	6	2.2	2101	5	PCT-US93-06160-4	Sequence 1, Appl
848	6	2.2	889	4	US-09-252-991A-30096	Sequence 30096, A	921	6	2.2	2142	4	US-09-540-236-459	Sequence 3459, Ap
849	6	2.2	891	4	US-09-585-858-4	Sequence 4, Appl	922	6	2.2	2175	4	US-09-404-650-2	Sequence 2, Appl
850	6	2.2	892	4	US-09-585-858-2	Sequence 2, Appl	923	6	2.2	2175	4	US-09-404-650-2	Sequence 2, Appl
851	6	2.2	892	4	US-09-585-858-3	Sequence 3, Appl	924	6	2.2	2188	4	US-09-404-650-4	Sequence 4, Appl
852	6	2.2	907	3	US-08-989-299-12	Sequence 12, Appl	925	6	2.2	2188	4	US-09-404-650-4	Sequence 4, Appl
853	6	2.2	907	3	US-09-407-427-12	Sequence 12, Appl	926	6	2.2	2314	4	US-09-258-541-4	Sequence 4, Appl
854	6	2.2	911	3	US-09-074-579-4	Sequence 4, Appl	927	6	2.2	2365	4	US-09-258-541-4	Sequence 4, Appl
855	6	2.2	911	3	US-09-388-774-4	Sequence 4, Appl	928	6	2.2	2431	4	US-09-343-681A-6304	Sequence 6304, Ap
856	6	2.2	917	4	US-09-489-039A-12582	Sequence 12582, A	929	6	2.2	2431	3	US-08-466-277-2	Sequence 2, Appl
857	6	2.2	939	5	PCT-US94-05905-20	Sequence 20, Appl	930	6	2.2	2464	4	US-09-376-594-726	Sequence 2, Appl
858	6	2.2	946	4	US-09-328-352-7973	Sequence 7973, Ap	931	6	2.2	2482	1	US-08-328-254-6	Sequence 6, Appl
859	6	2.2	948	4	US-09-613-303-21	Sequence 21, Appl	932	6	2.2	2595	3	US-09-370-700-2	Sequence 2, Appl
860	6	2.2	948	4	US-10-267-311-21	Sequence 21, Appl	933	6	2.2	2595	3	US-09-370-700-2	Sequence 2, Appl
861	6	2.2	951	4	US-09-924-097A-15	Sequence 15, Appl	934	6	2.2	2595	4	US-09-603-207-2	Sequence 2, Appl
862	6	2.2	957	4	US-09-252-991A-32568	Sequence 32568, A	935	6	2.2	2842	1	US-07-741-940-7	Sequence 7, Appl
863	6	2.2	980	4	US-09-252-991A-30838	Sequence 30838, A	936	6	2.2	2842	1	US-08-289-548A-7	Sequence 7, Appl
864	6	2.2	993	1	US-08-468-537-2	Sequence 2, Appl	937	6	2.2	2842	1	US-08-452-654-7	Sequence 7, Appl
865	6	2.2	1009	4	US-09-693-146-4	Sequence 4, Appl	938	6	2.2	2842	4	US-08-449-731-7	Sequence 7, Appl
866	6	2.2	1027	3	US-08-446-137B-2	Sequence 2, Appl	939	6	2.2	2843	1	US-07-741-940-2	Sequence 2, Appl
867	6	2.2	1032	4	US-09-976-59A-214	Sequence 214, App	940	6	2.2	2843	1	US-08-289-548A-2	Sequence 2, Appl
868	6	2.2	1075	4	US-09-198-452A-916	Sequence 916, App	941	6	2.2	2843	1	US-08-452-654-2	Sequence 2, Appl
869	6	2.2	1089	4	US-09-252-991A-20334	Sequence 20334, A	942	6	2.2	2843	1	US-08-452-655B-2	Sequence 2, Appl
870	6	2.2	1093	3	US-08-545-860D-55	Sequence 55, Appl	943	6	2.2	2843	1	US-08-452-655B-2	Sequence 2, Appl
871	6	2.2	1093	3	PCT-US94-04496-55	Sequence 55, Appl	944	6	2.2	2843	2	US-08-370-235A-2	Sequence 2, Appl
872	6	2.2	1102	4	US-09-358-383C-36	Sequence 36, Appl	945	6	2.2	2843	3	US-08-450-582-7	Sequence 2, Appl
873	6	2.2	1125	4	US-09-328-352-8233	Sequence 8233, Ap	946	6	2.2	2843	3	US-08-449-731-2	Sequence 7, Appl
874	6	2.2	1140	3	US-09-220-081-2	Sequence 2, Appl	947	6	2.2	2843	4	US-08-466-033-15	Sequence 15, Appl
875	6	2.2	1140	3	US-09-677-575-2	Sequence 2, Appl	948	6	2.2	2873	1	US-08-638-911A-2	Sequence 2, Appl
876	6	2.2	1156	4	US-09-198-452A-171	Sequence 171, App	949	6	2.2	2873	2	US-08-444-733-15	Sequence 15, Appl
877	6	2.2	1166	4	US-09-200-650E-7	Sequence 7, Appl	950	6	2.2	2873	2	US-08-444-733-15	Sequence 15, Appl
878	6	2.2	1196	4	US-08-680-328-35	Sequence 35, Appl	951	6	2.2	2873	2	US-08-461-361-15	Sequence 15, Appl
879	6	2.2	1228	4	US-09-439-313-537	Sequence 537, App	952	6	2.2	2873	2	US-08-461-361-15	Sequence 15, Appl
880	6	2.2	1228	4	US-09-636-215-537	Sequence 537, App	953	6	2.2	2873	2	US-08-461-361-15	Sequence 15, Appl
881	6	2.2	1228	4	US-09-685-166A-537	Sequence 537, App	954	6	2.2	2873	5	PCT-US95-06266-15	Sequence 15, Appl
882	6	2.2	1245	2	US-08-801-263A-3	Sequence 3, Appl	955	6	2.2	2905	4	US-08-469-260A-401	Sequence 401, App
883	6	2.2	1245	2	US-08-801-263A-6	Sequence 6, Appl	956	6	2.2	2905	4	US-08-467-344A-401	Sequence 401, App
884	6	2.2	1245	2	US-08-801-263A-10	Sequence 10, Appl	957	6	2.2	2905	4	US-08-467-344A-401	Sequence 401, App
885	6	2.2	1245	3	US-09-102-248-3	Sequence 3, Appl	958	6	2.2	2910	1	US-08-466-033-183	Sequence 183, App
886	6	2.2	1245	3	US-09-102-248-6	Sequence 6, Appl	959	6	2.2	2910	2	US-08-444-733-183	Sequence 183, App
887	6	2.2	1245	3	US-09-102-248-10	Sequence 10, Appl	960	6	2.2	2910	2	US-08-444-733-183	Sequence 183, App
888	6	2.2	1245	3	US-09-367-764-3	Sequence 3, Appl	961	6	2.2	2910	2	US-08-461-361-183	Sequence 183, App
889	6	2.2	1245	4	US-09-367-764-6	Sequence 6, Appl	962	6	2.2	2910	2	US-08-461-361-183	Sequence 183, App
890	6	2.2	1245	4	US-09-367-764-10	Sequence 10, Appl	963	6	2.2	2910	5	PCT-US95-06266-157	Sequence 157, App
891	6	2.2	1261	4	US-09-439-313-538	Sequence 313, App	964	6	2.2	2910	5	US-08-821-355A-7	Sequence 7, Appl
892	6	2.2	1261	4	US-09-439-313-538	Sequence 313, App	965	6	2.2	2910	5	US-09-003-687A-7	Sequence 7, Appl
893	6	2.2	1261	4	US-09-636-215-538	Sequence 538, App	966	6	2.2	2910	3	US-09-136-605-7	Sequence 7, Appl
894	6	2.2	1285	1	US-07-582-945-2	Sequence 2, Appl	967	6	2.2	3021	1	US-07-689-008-2	Sequence 2, Appl
895	6	2.2	1285	1	US-08-453-141-2	Sequence 2, Appl	968	6	2.2	3021	3	US-09-413-814-93	Sequence 93, Appl
896	6	2.2	1285	1	US-08-293-314-2	Sequence 2, Appl	969	6	2.2	3079	3	US-09-413-814-80	Sequence 80, Appl
897	6	2.2	1297	4	US-09-328-352-6373	Sequence 6373, Ap	970	6	2.2	3168	4	US-09-489-039A-14067	Sequence 14067, A
898	6	2.2	1341	4	US-09-252-991A-26785	Sequence 26785, A	971	6	2.2	3170	4	US-07-642-734C-5	Sequence 5, Appl
899	6	2.2	1374	4	US-09-252-991A-24636	Sequence 24636, A	972	6	2.2	3170	3	US-08-439-009A-5	Sequence 5, Appl
900	6	2.2	1427	4	US-09-252-991A-27005	Sequence 27005, A	973	6	2.2	3201	4	US-09-679-279-15	Sequence 15, Appl
901	6	2.2	1461	4	US-09-976-594-531	Sequence 531, App	974	6	2.2	3248	5	US-08-353-700-1	Sequence 1, Appl
902	6	2.2	1487	4	US-09-489-039A-12113	Sequence 12113, A	975	6	2.2	3348	5	PCT-US95-16216-1	Sequence 1, Appl
903	6	2.2	1646	4	US-09-535-008-67	Sequence 67, Appl	976	6	2.2	3491	2	US-07-642-734C-2	Sequence 2, Appl

```

977 6 2.2 3491 3 US-08-438-009A-2 Sequence 2, Appl
978 6 2.2 3546 4 US-09-679-279-13 Sequence 13, Appl
979 6 1.8 3898 4 US-08-750-717-2 Sequence 2, Appl
980 5 1.2 5 4 US-08-877-605-330 Sequence 330, Appl
981 5 1.8 6 1 US-08-704-170-81 Sequence 81, Appl
982 5 1.8 6 5 PCT-US94-02631-81 Sequence 81, Appl
983 5 1.8 7 1 US-08-346-333-83 Sequence 83, Appl
984 5 1.8 7 3 US-09-258-754-389 Sequence 389, App
985 5 1.8 7 3 US-09-042-107-389 Sequence 389, App
986 5 1.8 7 3 US-08-687-580-63 Sequence 63, Appl
987 5 1.8 7 3 US-09-457-046B-40 Sequence 40, Appl
988 5 1.8 7 4 US-09-731-558-19 Sequence 159, Appl
989 5 1.8 7 4 US-09-989-789-159 Sequence 159, App
990 5 1.8 7 4 US-09-989-789-275 Sequence 275, App
991 5 1.8 7 4 US-09-989-789-378 Sequence 378, App
992 5 1.8 7 4 US-09-989-789-383 Sequence 383, App
993 5 1.8 7 4 US-09-989-789-437 Sequence 437, App
994 5 1.8 7 4 US-09-989-789-633 Sequence 633, App
995 5 1.8 7 4 US-09-989-789-654 Sequence 654, App
996 5 1.8 7 4 US-09-989-789-760 Sequence 760, App
997 5 1.8 7 4 US-09-989-789-1077 Sequence 1077, App
998 5 1.8 7 4 US-09-989-789-1078 Sequence 1078, App
999 5 1.8 7 4 US-09-989-789-1761 Sequence 1761, App
1000 5 1.8 7 4 US-09-989-789-1765 Sequence 1765, App

```

ALIGNMENTS

```

RESULT 1
US-09-540-236-2991
; Sequence 2991, Application US/09540236
; Parent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2991
; LENGTH: 118
; TYPE: PRT
; ORGANISM: M.cattarhalis
US-09-540-236-2991

```

```

Query Match 20.7%; Score 57; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.2e-47; Indels 0; Gaps 0;
Matches 57; Conservative 0; Mismatches 0;

```

```

QY 47 PEOAAVEAAGVAKEXNLTVELEFNDYAMPNSAVSKGLDANAMQHRYLEKDSQ 103
DB 57 PEOAAVEAAGVAKEXNLTVELEFNDYAMPNSAVSKGLDANAMQHRYLEKDSQ 113

```

```

RESULT 2
US-09-543-681A-6534
; Sequence 6534, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; FILE REFERENCE: 2709.1003-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6534
; LENGTH: 279
; TYPE: PRT

```

```

; ORGANISM: Proteus mirabilis
US-09-543-681A-6534

```

```

Query Match 5.8%; Score 16; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2e-07; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0;

```

```

QY 224 FVEDKDSFYNNITVAR 239
DB 227 FVEDKDSFYNNITVAR 242

```

```

RESULT 3
US-09-328-352-6702
; Sequence 6702, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-033A
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6702
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6702

```

```

Query Match 3.6%; Score 10; DB 4; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.11; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

```

```

QY 195 ARAIDVDVLA 204
DB 200 ARAIDVDVLA 209

```

```

RESULT 4
US-09-489-039A-10641
; Sequence 10641, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10641
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10641

```

```

Query Match 3.3%; Score 9; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

```

```

QY 133 ELKDGAATIA 141
DB 140 ELKDGAATIA 148

```

```

RESULT 5
US-09-107-532A-7234
; Sequence 7234, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:

```

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: US/09/107,532A
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7234:
SEQUENCE CHARACTERISTICS:
LENGTH: 291 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...291
SEQUENCE DESCRIPTION: SEQ ID NO: 7234:
US-09-107-532A-7234

Query Match 3.3%; Score 9; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 PYNIIIVAR 239
|||
240 PYNIIIVAR 248

Db 240 PYNIIIVAR 248

RESULT 6
US-09-328-352-6924
Sequence 6924, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6924
LENGTH: 296
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6924

Query Match 3.3%; Score 9; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 NLRALIL 156
|||
164 NLRALIL 172

Db 164 NLRALIL 172

RESULT 7
US-09-107-532A-6443
Sequence 6443, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6443:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...269
SEQUENCE DESCRIPTION: SEQ ID NO: 6443:
US-09-107-532A-6443

Query Match 2.9%; Score 8; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 ELKDGAT 140
|||
118 ELKDGAT 125

Db 118 ELKDGAT 125

RESULT 8
US-09-252-991A-19629
Sequence 19629, Application US/09252991A
Patent No. 6551795

```

; GENERAL INFORMATION:
; APPLICANT: Marc U. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19629
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19629

Query Match      2.9%; Score 8; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 10;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      51 VAEVAGGV 58
      |||||
Db      98 VAEVAGGV 105

RESULT 9
US-09-543-681A-4548
; Sequence 4548, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4548
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4548

Query Match      2.9%; Score 8; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      122 AGYSTKIK 129
      |||||
Db      123 AGYSTKIK 130

RESULT 10
US-09-543-681A-6914
; Sequence 6914, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6914
; LENGTH: 279
; TYPE: PRT

; ORGANISM: Proteus mirabilis
US-09-543-681A-6914

Query Match      2.9%; Score 8; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      159 QGLIKKD 166
      |||||
Db      160 QGLIKKD 167

RESULT 11
US-09-489-039A-12337
; Sequence 12337, Application US/09489039A
; Patent No. 6610636
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12337
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12337

Query Match      2.9%; Score 8; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 VGNTEVVP 120
      |||||
Db      124 VGNTEVVP 131

RESULT 12
US-09-252-991A-17635
; Sequence 17635, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc U. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17635
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17635

Query Match      2.9%; Score 8; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 17;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      149 LARALILL 156
      |||||
Db      95 LARALILL 102

RESULT 13
```

```
US-08-858-207A-422
; Sequence 422, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328e1 Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmì, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 422:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
; US-08-858-207A-422

Query Match      2.5%; Score 7; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      101 DSOEKL 107
Db      34 DSOEKL 40

RESULT 14
US-09-134-001C-3331
; Sequence 3331, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3331
```

```
LENGTH: 116
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3331

Query Match      2.5%; Score 7; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      65 LTVELVE 71
Db      31 LTVELVE 37

RESULT 15
US-09-963-791-18
; Sequence 18, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399e1 Human Proteases and Polynucleotides Encoding the San
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-18

Query Match      2.5%; Score 7; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      233 VNIIVAR 239
Db      144 VNIIVAR 150

RESULT 16
US-08-572-447C-11
; Sequence 11, Application US/08572447C
; Patent No. 5955090
; GENERAL INFORMATION:
; APPLICANT: Knaap, Bernhard
; APPLICANT: Hungerer, Klaus-Dieter
; APPLICANT: Broker, Michael
; APPLICANT: Von Specht, Bernd-Ulrich
; APPLICANT: Domdey, Horst
; TITLE OF INVENTION: Immunogenic Hybrid Protein Ogrf-Opr1
; TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/572.447C
FILING DATE: 14-DEC-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4400
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-572-447C-11

Query Match 2.5%; Score 7; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 EVEAEAK 265
|||
155 EVEAEAK 161

Db 155 EVEAEAK 161

RESULT 17
US-09-267-747-11
Sequence 11, Application US/09267747
Patent No. 6300102
GENERAL INFORMATION:
APPLICANT: Knapp, Bernhard
APPLICANT: Hungerer, Klaus-Dieter
APPLICANT: Broker, Michael
APPLICANT: Von Specht, Bernd-Ulrich
APPLICANT: Domdey, Horst
TITLE OF INVENTION: Immunogenic Hybrid Protein Opr-Opr1
TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/267,747
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/572.447
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4400

TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-267-747-11

Query Match 2.5%; Score 7; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 EVEAEAK 265
|||
155 EVEAEAK 161

Db 155 EVEAEAK 161

RESULT 18
PCT-US94-02889-2
Sequence 2, Application PC/TUS9402889
GENERAL INFORMATION:
APPLICANT: Lefebvre Dr., Rance B.
APPLICANT: Perring Dr., Guey-Chuen
TITLE OF INVENTION: Methods, Compositions, and
TITLE OF INVENTION: Kits for Diagnosing Lyme
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBBINS, BERLINER & CARSON
STREET: 201 North Figueroa Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/02889
FILING DATE: herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Berliner, Robert
REGISTRATION NUMBER: 20,191
REFERENCE/DOCKET NUMBER: 5555-211-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 977-1001
TELEFAX: (213) 977-1003
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 199 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-02889-2

Query Match 2.5%; Score 7; DB 5; Length 199;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KYNLIVE 68
|||
16 KYNLIVE 22

Db 16 KYNLIVE 22

RESULT 19
US-08-572-447C-15
Sequence 15, Application US/08572447C
Patent No. 5555090

GENERAL INFORMATION:
APPLICANT: Knapp, Bernhard
APPLICANT: Hungerer, Klaus-Dieter
APPLICANT: Broker, Michael
APPLICANT: Von Specht, Bernd-Ulrich
APPLICANT: Domdey, Horst
TITLE OF INVENTION: Immunogenic Hybrid Protein Oprf-OprI
TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: US/08/572,447C
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4000
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-572-447C-15

Query Match 2.5%; Score 7; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 259 EVEAEAK 265
DB 220 EVEAEAK 226
RESULT 20
US-09-267-747-15
Sequence 15, Application US/09267747
Patent No. 6300102
GENERAL INFORMATION:
APPLICANT: Knapp, Bernhard
APPLICANT: Hungerer, Klaus-Dieter
APPLICANT: Broker, Michael
APPLICANT: Von Specht, Bernd-Ulrich
APPLICANT: Domdey, Horst
TITLE OF INVENTION: Immunogenic Hybrid Protein Oprf-OprI
TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: US/09/267,747
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/572,447
FILING DATE: 14-DEC-1995
APPLICATION NUMBER: EP 94120023.0
FILING DATE: 16-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: M. Paul Barker
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 05552.1395-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4000
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-267-747-15

Query Match 2.5%; Score 7; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 259 EVEAEAK 265
DB 220 EVEAEAK 226

RESULT 21
US-09-543-681A-4708
Sequence 4708, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 4708
LENGTH: 274
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-4708

Query Match 2.5%; Score 7; DB 4; Length 274;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 149 IAPALIL 155
DB 166 IAPALIL 172

RESULT 22
US-09-489-039A-7722
Sequence 7722, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:

```
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7722
LENGTH: 275
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7722

Query Match      2.5%; Score 7; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      149 LARATL 155
DB      167 LARATL 173

RESULT 23
US-09-434-354-48
Sequence 48; Application US/09434354
Patent No. 6562563
GENERAL INFORMATION:
APPLICANT: Murphy, Anne N.
APPLICANT: Clevenger, William
APPLICANT: Wiley, Sandra Eileen
APPLICANT: Andreyev, Alexander Y.
APPLICANT: Frigeri, Luciano G.
APPLICANT: Velicelcib, Gonul
APPLICANT: Davis, Robert E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETERMINING
TITLE OF INVENTION: INTERACTIONS OF MITOCHONDRIAL COMPONENTS, AND FOR
FILE REFERENCE: 660088.433
CURRENT APPLICATION NUMBER: US/09/434,354
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 54
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 48
LENGTH: 298
TYPE: PRT
ORGANISM: Homo sapien
US-09-434-354-48

Query Match      2.5%; Score 7; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      29 AAISKTA 35
DB      19 AAISKTA 25

RESULT 24
US-09-434-354-49
Sequence 49; Application US/09434354
Patent No. 6562563
GENERAL INFORMATION:
APPLICANT: Murphy, Anne N.
APPLICANT: Clevenger, William
APPLICANT: Wiley, Sandra Eileen
APPLICANT: Andreyev, Alexander Y.
APPLICANT: Frigeri, Luciano G.
APPLICANT: Velicelcib, Gonul
APPLICANT: Davis, Robert E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETERMINING
TITLE OF INVENTION: INTERACTIONS OF MITOCHONDRIAL COMPONENTS, AND FOR
```

```
FILE REFERENCE: 660088.433
CURRENT APPLICATION NUMBER: US/09/434,354
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 54
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 49
LENGTH: 298
TYPE: PRT
ORGANISM: Homo sapien
US-09-434-354-49

Query Match      2.5%; Score 7; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      29 AAISKTA 35
DB      19 AAISKTA 25

RESULT 25
US-09-134-001C-3171
Sequence 3171; Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3171
LENGTH: 302
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3171

Query Match      2.5%; Score 7; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      182 KKVYKE 188
DB      3 KKVYKE 9

RESULT 26
US-09-252-991A-18873
Sequence 18873; Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18873
LENGTH: 306
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18873
```

Query Match 2.5%; Score 7; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 229 DSPYNI 235
Db 259 DSPYNI 265

RESULT 27
US-09-522-714-10
; Sequence 10, Application US/09522714
; Patent No. 6563020
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Yalpani, Nasser
; TITLE OF INVENTION: Maize Chitinases and Their Use in
; FILE REFERENCE: 1100
; CURRENT APPLICATION NUMBER: US/09/522,714
; EARLIER FILING DATE: 2000-03-10
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Zea mays
US-09-522-714-10

Query Match 2.5%; Score 7; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 VGLTASE 220
Db 254 VGLTASE 260

RESULT 28
US-09-963-791-8
; Sequence 8, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-8

Query Match 2.5%; Score 7; DB 4; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 VNIIVAR 239
Db 295 VNIIVAR 301

RESULT 29
US-09-963-791-16
; Sequence 16, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the San
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-16

Query Match 2.5%; Score 7; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 VNIIVAR 239
Db 144 VNIIVAR 150

RESULT 30
US-09-489-039A-9512
; Sequence 9512, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9512
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9512

Query Match 2.5%; Score 7; DB 4; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 DIVENPK 182
Db 222 DIVENPK 228

RESULT 31
US-09-540-236-2111
; Sequence 2111, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2111
LENGTH: 342
TYPE: PRT
ORGANISM: M.catarrhalls
US-09-540-236-2111

Query Match 2.5%; Score 7; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 LARALIL 155
DB 1 LARALIL 7

RESULT 32
US-09-543-681A-5961
Sequence 5961, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETTON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 5961
LENGTH: 347
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-5961

Query Match 2.5%; Score 7; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 LARALIL 155
DB 148 LARALIL 154

RESULT 33
US-09-252-991A-30094
Sequence 30094, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30094
LENGTH: 351
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30094

Query Match 2.5%; Score 7; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 EVEAEAK 265
DB 345 EVEAEAK 351

RESULT 34
US-09-963-791-20
Sequence 20, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399 Human Proteases and Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20
LENGTH: 356
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-20

Query Match 2.5%; Score 7; DB 4; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 VNIIVAR 239
DB 144 VNIIVAR 150

RESULT 35
US-09-543-681A-5390
Sequence 5390, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETTON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 5390
LENGTH: 361
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-5390

Query Match 2.5%; Score 7; DB 4; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 EAEAKQ 267
DB 174 EAEAKQ 180

RESULT 36
US-08-305-172B-4
Sequence 4, Application US/08305172B
Patent No. 5656470

GENERAL INFORMATION:
APPLICANT: Martinis, Susan A.
APPLICANT: Zhang, Jiansu
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Mycobacterial Seryl-tRNA
TITLE OF INVENTION: Synthetase Genes, Tester Strains and Assays
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Willetta Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/305,172B
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-305-172B-4

Query Match 2.5%; Score 7; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 258 DEVEAEA 264
Db 87 DEVEAEA 93

RESULT 37
US-09-963-791-22
Sequence 22, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 438
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-22

Query Match 2.5%; Score 7; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 233 VNIIVAR 239
Db 144 VNIIVAR 150

RESULT 38
US-09-252-991A-20116
Sequence 20116, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20116
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20116

Query Match 2.5%; Score 7; DB 4; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LASGIAL 18
Db 1 LASGIAL 7

RESULT 39
US-09-963-791-6
Sequence 6, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 468
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-6

Query Match 2.5%; Score 7; DB 4; Length 468;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 VNIIVAR 239
Db 295 VNIIVAR 301

RESULT 40

US-07-686-591-4
; Sequence 4, Application US/07686591
; Patent No. 5215915
; GENERAL INFORMATION:
; APPLICANT: Tiberi, Mario
; APPLICANT: Jarvie, Keith R.
; APPLICANT: Caron, Marc G.
; TITLE OF INVENTION: Cloned Gene Encoding Rat D18 Dopamine Receptor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5215915th Carolina
; COUNTRY: U.S.A.
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/686,591
; FILING DATE: 19910406
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-686-591-4

Query March 2.5%; Score 7; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AVAEVAG 56
|||
Db 97 AVAEVAG 103

Search completed: June 16, 2004, 11:18:09
Job time : 31 secs